SYNTHESIS AND CHARACTERIZATION OF PHOSPHONO-CHEY FROM *THERMOTOGA MARITIMA*

R. Matthew Haas

A Thesis Submitted to the
University of North Carolina Wilmington in Partial Fulfillment
Of the Requirements for the Degree of
Master of Science

Department of Chemistry and Biochemistry

University of North Carolina Wilmington

2007

Approved by

Advisory Committee

Paulo Almeida Sridhar Varadarajan

Christopher Halkides
Chair

Accepted by

Robert Roer
Dean, Graduate School

This thesis has been prepared in the style and format

Consistent with the journal Biochemistry

TABLE OF CONTENTS

ABSTRACT

The chemical and molecular mechanisms of bacterial chemotaxis have been extensively characterized through structural studies of mutant chemotaxis proteins. CheY~P (the active, signaling form of CheY) binds the switch protein FliM at the intracellular base of the flagellar motor, switching the direction of flagellar rotation from clockwise to counterclockwise and punctuating periods of smooth-swimming with tumbling and random reorientation. The conformational changes accompanying phosphorylation and activation of the response regulator CheY have been identified; however, transfer of the "active" signal from CheY~P to the flagellar switch complex is unknown. Crystallization of FliM alone and in complex with CheY~P would allow detailed analysis of the interactions between CheY~P and FliM that force a change in the directional rotation of the flagella. These structures are difficult to crystallize since FliM is a membrane-bound protein and is not soluble in the conditions necessary for crystallization. The α-proteobacteria, *Bacillus subtilis* and *Thermotoga maritima* contain protein phosphatases that share homology with the central domain of FliM. If these phosphatases bind CheY~P in a similar manner as FliM, a crystal structure of CheY~P in complex with one of these phosphatases may shed light on the interactions that occur at the base of the flagellar motor. Since CheY~P has a short half life, an analogue known as phosphono-CheY was synthesized and purified under various conditions. *Thermotoga maritima* D54C/C81S CheY was purified from the soluble fraction of a cell lysate using immobilized metal affinity chromatography (Ni^{2+}-NTA) and size-exclusion chromatography. Purified CheY was reacted with phosphonomethyltriflate under alkaline conditions in the presence of a divalent (or trivalent) metal yielding roughly 50% phosphono-CheY by RP-HPLC and reaction with Ellman's reagent. Phosphono-CheY was separated from un-reacted CheY by first labeling the unmodified protein with biotin and then

isolating the two species by avidin affinity chromatography or by hydrophobic interaction chromatography. The presence of phosphono-CheY was confirmed by electrospray ionization mass spectrometry of the intact protein. Mass spectrometry also suggested the presence of multiply-phosphonomethylated protein and oxidized protein within a phosphonomethylation reaction mixture and experiments to disprove this hypothesis are on-going.

ACKNOWLEDGEMENTS

First and foremost I want to thank Cory Bottone for the endless discussions we shared on CheY. I would like to thank Andy Alpert of PolyLC and Amos Heckendorf of The Nest Group for their unending support and extensive knowledgebase on HPLC. I would like to thank Will Boomershine for his willingness to listen and answer even the simplest of questions, and for his assistance in mass spectrometry. I sincerely thank Thomas Schuster and George Dubay for all the work they put in to obtain the mass spec data, and I must acknowledge Will Thompson at GSK for his advice on interpreting protein mass spectrum. I would like to thank Rick Dahlquist and Will Deutschman for their insights on Thermotoga proteins. I would also like to sincerely thank our collaborators Brian Crane and Gabriela Gonzales-Bonet for offering our lab and I the opportunity to work on such a great project and for providing both protein and the plasmid DNA for *T. maritima* D54C/C81S CheY. Finally, I would like to thank my advisor, Dr. Christopher Halkides, who generously gave me the opportunity to work as a research assistant on such a challenging and personally rewarding project.

DEDICATION

This thesis is for everyone who believed in me. Mom, Dad, Aunt Ann, Melinda, Cory, D, Jay, Liz, Sonia, Lindsay, Jaclyn, Adam, Josh, Charrise, Dr. Varadarajan, Dr. Almeida, and Dr. Halkides; thank you for you advice and support.

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Signal Transduction

Survival, particularly for single-celled organisms, is directly related to the ability to sense changes in a dynamic environment and react accordingly. Cells efficiently regulate processes such as development, sporulation, virulence, nutrient acquisition, osmoregulation, and chemotaxis by monitoring important changes in environmental conditions (*1*). For example, unless the surroundings are harsh, sporulating bacteria do not dedicate the resources for spore coat synthesis. Additionally, only when the cell is exposed to a lactose-rich environment will it increase intracellular production of β-galactosidase (an enzyme catalyzing lactose degradation), otherwise it would be a waste of energy and materials. For this reason, cells have multiple receptors on their surface that enable them to sense a wide variety of environmental signals. These receptors are linked to the inside of the cell through transmembrane domains, so the information gathered from the surroundings can be transformed into a usable intracellular form. Specialized proteins within the cell, collectively called signaling proteins, are responsible for interpreting the various signals and coordinating an appropriate response. The cell can therefore effectively control how it responds to extracellular stimuli by controlling the activity of its signaling proteins.

The activity of signaling proteins is regulated both covalently and non-covalently. Two well-known models describe enzyme regulation: the allosteric model and ATP-dependent phosphorylation and dephosphorylation, both of which involve modifications that alter the structure of the protein. Allosteric regulation involves the binding of a small molecule or allosteric modulator to a regulatory domain separate from the active site of the protein. Conformational changes occur upon binding that alter the structure of the active state and

consequently the activity of the enzyme. In *Escherichia coli* for example, aspartate transcarbamoylase is switched off by the binding of the allosteric modulator cytosine triphosphate to the enzyme's regulatory domains (*2*). This causes the regulatory domains to pull the catalytic domains close together, hiding the active site cleft from any substrate. Enzymatic activity is also regulated by phosphorylation and dephosphorylation, in which protein kinases catalyze the transfer of the γ-phosphoryl group of ATP to a residue often in the enzyme's active site. The phosphoryl group is not a large moeity, but it can have important effects within the structure of proteins. First, the two negatively charged oxygens attract regions of positively charged residues that can alter the binding surface of the protein. Second, phosphoryl groups introduce two additional sites for hydrogen bonding or salt bridge formation between amino acid side chains. Third, a phosphoryl group present in the active site can be part of a structure directly recognized by the binding sites of other proteins. A well-understood example of this is the SH2 domains in eukaryotes, which recognizes and binds short peptide sequences containing phosphorylated tyrosine residues (*2*). Without the presence of the phosphoryl group, no binding occurs.

A cell's signaling proteins are controlled both allosterically and by phosphorylation, and they essentially adopt one of two conformations: a conformation associated with the active state of the enzyme and a conformation associated with the inactive state of the enzyme. In its active conformation the protein can bind and either activate or inactivate the next protein or set of proteins in the pathway. In eukaryotes for example, signaling proteins are often protein kinases organized into a phosphorylation cascade. One protein kinase, activated by phosphorylation, catalyzes the transfer of a phosphoryl group to the next protein kinase and so on. The phosphorylation cascade is terminated by phosphatases that catalyze the removal of the

phosphoryl group, switching the proteins back to their inactive state. In eukaryotes, these signaling pathways are often complex webs of interactions between proteins and other small molecule mediators in which one enzyme will amplify the signal by activating more than one downstream signaling molecule. Much simpler signaling systems control the way prokaryotes respond to environmental stimuli. These systems, collectively termed two-component systems, are found only in prokaryotes and certain lower eukaryotes.

Two-Component Signal Transduction

In two-component signal transduction, an autohistidine kinase (HK) receives a signal from transmembrane receptors and reversibly transfers a phosphoryl group to a conserved aspartate residue at the active site of the response regulator (RR), controlling its signaling state. The HK generally contains an N-terminal transmembrane domain that directly binds extracellular ligands and a C-terminal kinase domain. The RR generally contains an N-terminal phosphoryl-accepting regulatory domain and a C-terminal effector domain responsible for enzymatic activity or binding target genes. The proteins in these systems are comprised of highly conserved domains (*3*) that are adaptable and can be arranged in different combinations to suit the specific needs of individual systems (*4*). The modularity of these domains allows multiple systems to interact with one another, providing a cell with the ability to coordinate a concerted response to multiple, sometimes conflicting stimuli.

A well-studied example of the two-component signal transduction system is that regulating the osmotic pressure of the intracellular environment of bacteria. Osmoregulation is controlled by the two-component system in which the histidine kinase EnvZ transfers a phosphoryl group to an aspartate within the N-terminal domain of the response regulator OmpR.

Activation of OmpR by phosphorylation is accompanied by a conformational change that enhances the DNA-binding affinity of its C-terminal domain and regulates the transcription of the genes encoding the major outer membrane porin proteins *ompC* and *ompF*. The C-terminal domain of OmpR also interacts with the alpha subunit of RNA polymerase to efficiently regulate transcription. The majority of response regulators are similar to OmpR and contain two or more domains; however, the most extensively characterized response regulator is the single-domain CheY protein.

Bacterial Chemotaxis

CheY is a single domain response regulator that functions in the chemotaxis response (*5*). Once the cytoplasmic (no transmembrane domain) histidine kinase CheA transfers a phosphoryl group from its catalytic histidine residue to the active site aspartate (57 in *E. coli;* 54 in *T. maritima*) of CheY (*6*), the phosphorylated state of CheY, CheY~P, binds to the switch protein FliM at the base of the flagellar motor and changes its direction of rotation from counterclockwise, generating smooth-swimming behavior, to clockwise, generating tumbly behavior (*7*). Changes in the concentration of CheY~P determine how frequently the periods of smooth-swimming are punctuated by tumbles; these changes can create a biased-random walk toward a better chemical environment. Mutations in CheY that bring about a greater tendency either toward smooth-swimming or toward tumbly behavior impair chemotaxis relative to wild-type (*8*).

Chemotaxis is both a robust and adaptive pathway. Bacteria are able to sense temporal changes in solute concentrations of a few parts per thousand over background levels ranging 5 orders of magnitude (*9, 10*). The cellular level of CheY~P must be carefully controlled in order

for the microorganism to translate such small changes in the extracellular environment to large alterations in swimming behavior. This is accomplished by adaptation pathways at the receptors and by dephosphorylation of the response regulator, a process catalyzed by its own autophosphatase activity as well as by a variety of phosphatases (CheZ in *E. coli*, CheX, CheC/CheD, and FliY/N in *T. maritima*).

Adaptation is accomplished by the complementary behavior of two other response regulators, CheR and CheB. CheR is a methyl esterase that transfers methyl groups to conserved glutamine residues exposed on the intracellular portion of transmembrane receptors. CheB is a methyl transferase that demethylates these glutamine residues. The rate at which CheA is phosphorylated (and hence the level of CheY~P) depends on the methylation state of the receptor. The cell is therefore able to decrease or increase the level of its response to extracellular stimuli based on this level of methylation. For example, when the cell finds itself in an environment with a consistently high concentration of sucrose, its receptors will adapt to treat the high level of sucrose as if it were normal. This ensures that CheY is not over-phosphorylated and overactive, but that it remains sensitive to small temporal changes in the extracellular environment. Rapid dephosphorylation of CheY is another mechanism that prevents overactive signaling at the flagellar motor. This occurs by the activity of CheY-specific phosphatases (*11*) as well as by intrinsic autodephosphorylation of CheY (*12*).

Structure/Function Studies of CheY

A number of structures of *E. coli* CheY in both its active and inactive states exist in the literature (*13-20*). It has $(\beta/\alpha)_5$ saddle topology with its active site near the C-terminus of the central β-sheet (Figure 1). Five highly conserved residues are found near the active site: Asp12,

Asp13, Asp57, Thr87, and Lys109. Mutation at any of these conserved residues impairs chemotaxis (*21-24*). The third conserved aspartate, Asp57, is the site of phosphorylation (*25*). Mutation of this residue produces a protein that cannot be phosphorylated at the active site and does not bind FliM (*22, 26, 27*). Thr87 is involved in the hydrogen bond network at the active site through its γ-hydroxyl group. Certain mutations at residue 87 produce CheY proteins that can be phosphorylated but non-signaling (*23, 28*). This suggests the importance of residue 87 in propagating conformational changes in the protein upon phosphorylation. T87I CheY is slower in both its autophosphatase activity and its CheZ-promoted phosphatase activity (*23*). The side-chain of Lys109 binds to the phosphoryl group of residue 57 (*18*). Mutant K109R is highly phosphorylatable *in vitro* but exclusively non-signaling *in vivo* (*26*). The high level of phosphorylation seen in K109R CheY was shown to be due to lack of responsiveness to the phosphatase CheZ as well as a decrease in autophosphatase activity (*22*). Similarly to Thr87, Lys109 must be involved in a post-phosphorylation event.

Residue 106 is also a conserved amino acid. It is present as an aromatic residue (tyrosine or phenylalanine) in 80% of known response regulators (*3*). Matsumura and collaborators (*24*) confirmed that an aromatic amino acid at position 106 is required for proper CheY~P signaling. Mutagenesis and structure-function studies indicate that both the identity and the rotameric position of residue 106 are important for CheY~P signaling (*8, 23, 24*). Substitution of Tyr106 in *E. coli* CheY with tryptophan (Y106W CheY) produces a phosphorylation-dependent hyperactive mutant that generates mainly clockwise rotational bias in the presence of a phosphoryl-group donor. Replacement of Tyr106 with a nonaromatic, nonpolar residue results in completely smooth-swimming cells that are non-chemotactic.

In crystals of wild-type apo-CheY, Tyr106 is found at the FliM binding surface of CheY (*15, 29*) and two rotameric conformations are evident from the electron density: an inside, solvent-inaccessible position and an outside, solvent-exposed position. Crystals of Y106W CheY show that tryptophan is found exclusively as the solvent-inaccessible rotamer (*8*). In the structure of T87I CheY, Tyr106 is found only as the solvent-accessible rotamer (*23*). The addition of the ethyl moiety at residue 87 sterically blocks residue 106 from occupying the hydrophobic cavity where it is seen in activated structures. Combining these two mutations in T87I/Y106W CheY gives the same phenotype as T87I CheY (*8*). This result suggests that the buried rotameric conformation of residue 106 is required for CheY~P to induce clockwise rotation of the flagellar motor.

Rationale for the Study

Studies of the structure and function of the chemotaxis proteins have identified the residues important for CheY~P association with FliM (*29-31*), CheZ (*32, 33*), and CheA (*34, 35*). The mechanism of phosphoryl transfer from CheA to CheY has been extensively studied through crystallization (*34, 36-38*), and to a lesser extent the CheZ-promoted hydrolysis of PO_3-Asp-CheY (*11*). FliM, however, remains the most elusive protein of which to obtain a crystal structure, and characterization of the FliM-CheY~P complex has focused mainly on the binding of CheY~P to short, 16 residue peptides derived from FliM. Crystal structures of FliM and the FliM-CheY~P complex as a whole would explain the conformational changes that accompany signal transfer from CheY~P to FliM.

FliM proteins aggregate near the plasma membrane at the base of the flagellar motor forming a ring of ~34 FliM subunits. Membrane proteins such as this are difficult to crystallize.

The protein phosphatases CheX, CheC, and FliY/N of the α-proteobacteria *Bacillus subtilis* and *Thermotoga maritima* have sequence homology with the central domain of FliM (*39*). If these phosphatases are similar to FliM in the manner in which they interact with CheY~P, crystallization of the CheX-CheY~P complex or the CheC-CheY~P complex may better represent the FliM-CheY~P complex than just CheY~P bound to the FliM peptide alone. *E. coli* does not contain these phosphatases so *Thermotoga maritima* CheY was selected for this study.

A non-hydrolyzable analogue of CheY~P must be created if crystallography of the active form of the protein is to be done, since the half-life of the wild-type *E. coli* CheY~P is limited to less than a second by both its own autophosphatase activity as well as dephosphorylation by CheZ (*40*). Phosphono-CheY (*41*) is a crystallizable analogue of CheY~P that resembles CheY~P in the manner in which it binds the CheZ and FliM peptides (*42*). The goal of this project is to synthesize and purify phosphono-CheY from *Thermotoga maritima* D54C/C81S CheY. The protein will be used in crystallization trials of complexes with the CheX and CheC phosphatases, as well as FliM cross-linking experiments and isothermal titration calorimetry to measure binding of the analogue to the flagellar switch protein (*43*).

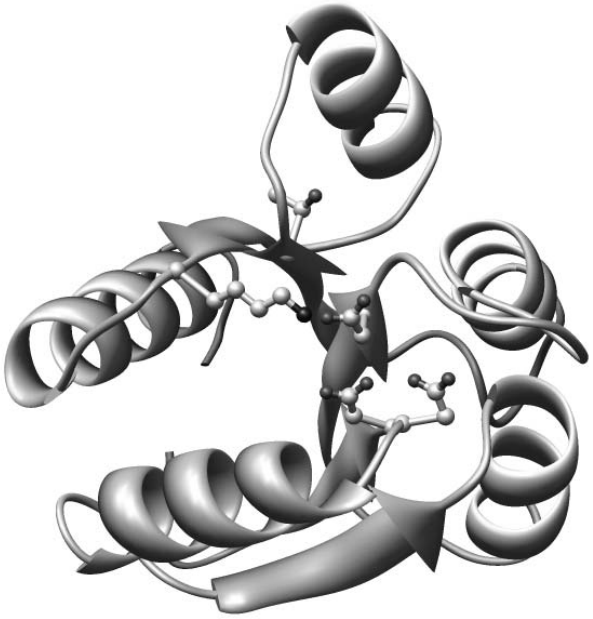


Figure 1. Three dimensional structure of *E. coli* CheY, a protein with $(\beta/\alpha)_5$ saddle topology and its active site near the C-terminus of the central β-sheet . Active site residues Asp12, Asp13, Asp57, Thr87, and Lys109 are displayed in ball and stick and colored by atom (N=blue, O=red).

MATERIALS AND METHODS

High Performance Liquid Chromatography

HPLC was the standard analytical method for determining the purity of protein prepared in the lab as well as to analyze the phosphonomethylation reactions and its subsequent purification. An HP1050 LC machine equipped with a quaternary pump, degasser, column oven, and variable wavelength detector was the primary instrument and an HP1100 equipped with a quaternary pump, degasser, column oven, and diode array detector was the secondary instrument used for high performance analytical and preparative chromatography. Both instruments were equipped with a pre-column filter that housed a 2 micron frit. Two Vydac 300 Å pore size C-18 columns with dimensions 2.1 mmID x 250 mm and 4.6 mmID x 250 mm (Nest Group) were used in reversed phase mode with water/acetonitrile (ACN) mobile phases and trifluoroacetic acid (TFA) as an ion-pairing agent. The appropriate guard column was placed upstream to minimize binding of contaminants to the column. Mobile phase A consisted of 95% water, 5% ACN, and 0.1% TFA and mobile phase B consisted of 20% water, 80% ACN, and 0.08% TFA. Flow-rates were 200 µL/min for the 2.1 mmID column and 1.00 mL/min for the 4.6 mmID column with pressures around 115 bar. *E. coli* D57C CheY eluted at 50.2% B (42.7% ACN) on a shallow gradient of 45.5% B to 52.5% B over 45 minutes, and *T. maritima* D54C/C81S eluted at 53.2% B (44.9% ACN) on a shallow gradient of 46.5% B to 54.0% B over 50 minutes. Occasionally, a second peak eluting at 42.1% ACN appeared in the RP-HPLC analysis of *T. maritima* D54C/C81S CheY. This peak was unknown and is herein referred to as CheY*. In the analysis of both *E. coli* and *T. maritima* phosphonomethylations, phosphono-CheY appeared between 12-16 minutes prior to unmodified CheY, which eluted near the end of the gradient (Figure 2). Samples for injection were prepared with 15 µg of protein (for the 2.1 mm ID

column) and 45 µg of protein (for the 4.6 mm ID column) brought up to 60 µL in mobile phase with a final acetonitrile concentration 5% below that of the starting conditions. Protein concentrations were determined by Bradford assay or absorbance at 280 nm prior to preparing the HPLC sample. A complete-filling technique utilizing a 60 µL injection into a 20 µL sample loop was used to maximize the reproducibility between injections.

A second type of HPLC termed hydrophobic interaction chromatography was used as a preparative method and rarely as an analytical method (*44*). A Polypropyl Aspartamide column (PolyLC) with a 1500 Å pore size and dimensions 4.6 mm ID x 150 mm was used. Mobile phase A was 1.6 M ammonium sulfate in 20 mM potassium phosphate, pH 6.5 and mobile phase B was 20 mM potassium phosphate, pH 6.5. The column was conditioned according to manufacturer's directions and was always stored in the low salt buffer or water. A gradient of 0 - 90% A was sufficient to separate phosphono-CheY from biotinylated CheY (PEO-Iodoacetyl). The best separation between phosphono-CheY and unmodified CheY was achieved with a three-step gradient: 5 minutes of 0 - 2% B, 40 minutes of 2 - 4% B, and 5 minutes of 4 - 6% A.

A third type of HPLC used a weak anion-exchange column (*45*). A PolyWAX LP column (PolyLC) of 4.6 x 200 mm was used with 5 mM Tris-Acetate, pH 7.9 and a gradient of 0 – 300 mM sodium acetate. The flow rate was 1.00 mL/min and the pressure around 100 bar. The column was conditioned according to manufacturer's instructions and an overnight (16.5 hrs) passivation with 40 mM EDTA·2Na at a flow-rate of 0.10 mL/min was used to scavenge any heavy metals in the system. Approximately four 30 min gradients of 0 – 300 mM NaCl and four 30 min gradients of 0 – 1 M NaCl were required to rid the system of residual EDTA before analyses were performed. *T. maritima* D54C/C81S CheY was buffer exchanged into 5 mM Tris-Acetate, pH 7.9 (mobile phase A) to a concentration of 4.4 mg/mL prior to injection and

phosphono-CheY collected from RP-HPLC and lyophilized to powder was re-suspended directly into 5 mM Tris-Acetate, pH 7.9 to a concentration of 2.3 mg/mL. 30 µL injections were made into a 20 µL sample loop. The runs consisted of a 20 min isocratic elution with 100% MP A followed by a 30 min gradient to 1.0 M NaCl and a hold for 20 min at 1.0 M NaCl. Peaks throughout the run were collected for analysis by Bradford assay and RP-HPLC.

Production and Purification of *E. coli* D57C CheY

E. coli D57C CheY cloned into a pCW plasmid was from the lab of Frederick Dahlquist. The plasmid was transformed into *E. coli* cell line Ecoli-B (Carolina Biological Supply Co.) by standard electroporation methods (Appendix A). Selection with 100 µg/mL ampicillin was used to identify colonies that had incorporated the plasmid, and test growths were performed to demonstrate proper IPTG induction before the transformed cell line was frozen with 2x freezing medium.

Two 5 mL overnight growths in Lauria broth containing 100 µg/mL ampicillin were inoculated with *E. coli* cell line Ecoli-B transformed with the D57C plasmid. The growths were incubated at 37 °C overnight shaking at ≥ 200 RPM. Using sterile technique, two 1 L growths in Lauria broth containing 100 µg/mL ampicillin were inoculated with 1 mL from the overnight cultures and incubated at 37 °C shaking at ≥ 200 RPM. The growths were monitored by the absorbance at 600 nm until the optical density reached 0.5, at which point protein production was induced with 1.0 mM IPTG. The flasks were kept shaking at 37 °C until stationary phase was reached. The cells were harvested at 4,500 RPM for 20 minutes at 4 °C in a Beckman JA-10 rotor. The supernatant was poured off and the cell pellets were gently resuspended at 0.2 g wet cell weight per 1 mL of sonnication buffer. 1.0 mM PMSF and 10 mM 2-mercaptoethanol (2-

ME) were added. Occasionally lysozyme was added to a final concentration of 0.2 mg/mL to assist in the breaking of cells.

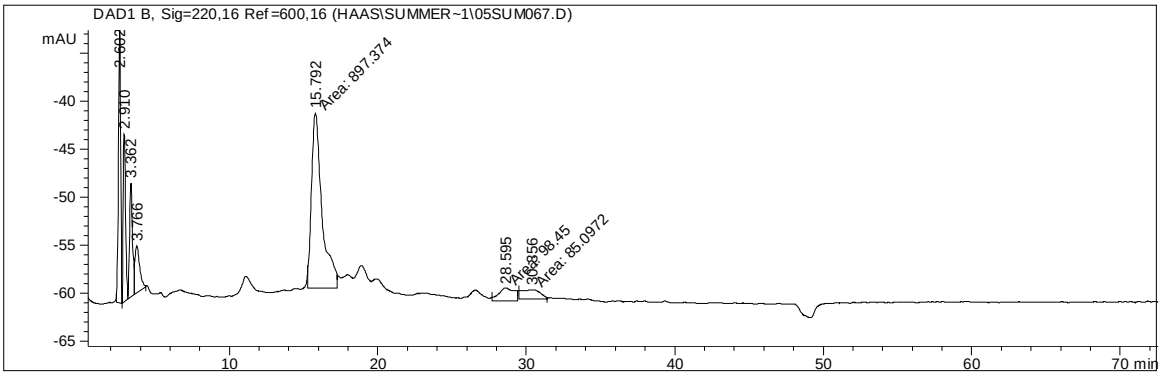
Sonnication was performed with a Fisher 550 Sonic Dismembrator with clean polished probes, re-tuning as necessary. 30 s bursts near power setting 7 were used and the suspension chilled to between 2 – 3 °C between each burst keeping care not to cause foaming of the solution. Cell lysis was checked by comparing the sonnicated cells to a small amount of unsonnicated cells under oil immersion light microscopy at 1000x magnification. Once it was determined that the majority of cells had been broken, the cell debris was spun at 17,000 RPM in a JA-20 rotor for 50 min at 4 °C. The insoluble fraction (sonnication pellet) was identified by SDS-PAGE to contain the greatest proportion of D57C CheY and was purified.

The sonnication pellet was resuspended in 30 mL of sucrose/EDTA (Appendix B) solution using a Dounce tissue homogenizer. This was spun at 17,000 RPM in a JA-20 rotor for 60 min at 4 °C. The supernatant was saved and the pellet resuspended in 30 mL of Triton/Tris buffer (Appendix B). This was spun at 16,000 RPM in a JA-20 rotor for 45 min at 4 °C. The supernatant was again saved and the pellet gently resuspended in 20 – 30 mL of Tris/Urea (Appendix B) and allowed to incubate overnight to allow the urea to dissolve the protein. This was spun at 17,000 RPM for 45 min in a JA-20 rotor at 4 °C and the supernatant saved. The pellet was again gently resuspended in Tris/Urea, allowed to incubate overnight, and spun down again. SDS-PAGE was used to check all of the supernatants for the presence of CheY.

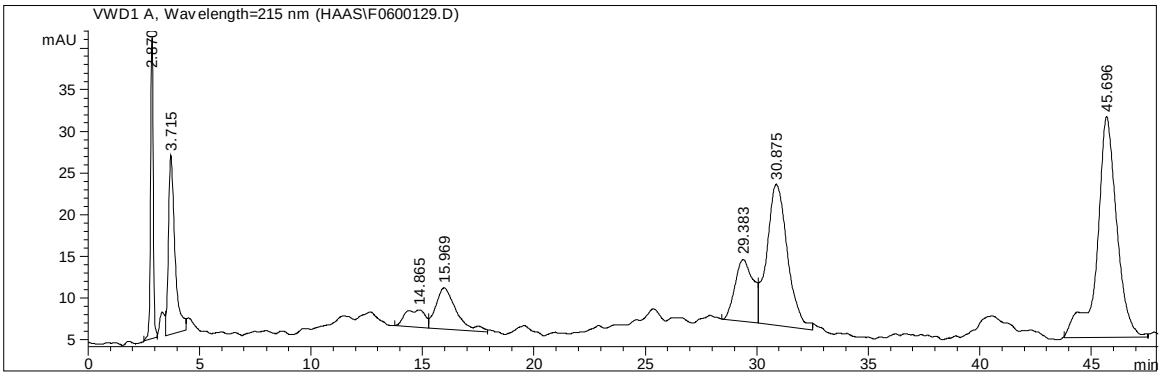
Purification of the urea supernatant proceeded by diluting the protein to 0.3 mg/mL in 50 mM Tris, pH 7.51 with 1 mM EDTA and a final urea concentration of 6 M. 10 mM 2-ME and 1 mM PMSF were also included. The protein was then dialyzed 3-4 times against 5 mM Tris, pH 7.51 with 1 mM EDTA and without urea, again including 2-ME and PMSF. A 50 mL (2.5 x 10

cm) DE-52 anion exchange column in the chloride form was equilibrated with the same Tris buffer and the conductivity of the column measured to ensure that all of the sodium chloride had been rinsed away. The protein is then loaded onto the column at 1 mL/min assuming its conductivity is near the conductivity of the starting buffer (i.e. the dialysis was complete). The column is then rinsed with several column volumes of Tris buffer until the absorbance at 280 nm has fallen to near that of the starting buffer. A 1 L gradient of 0-300 mM NaCl was passed through the column and 10 mL fractions collected at a flow-rate of 1 mL/min and the fractions analyzed by absorbance at 280 nm and SDS-PAGE for the presence of D57C CheY. The protein was then pooled based on purity and dialyzed 3-4 times into cibacron blue buffer, pH 7.9 with 2-ME (Appendix B).

A 50 mL cibacron blue column was equilibrated in cibacron blue buffer and the protein was loaded onto the column at 1 mL/min. The column was then rinsed with cibacron blue buffer and the protein eluted with 2 M NaCl in cibacron blue buffer. Fractions were collected and the protein combined based on absorbance at 280 nm and SDS-PAGE. The protein was then dialyzed into 20 mM BES, pH 7.01, with 1 mM EDTA, 0.02 % azide, and 10 mM DTT. The protein was concentrated to approximately 1 mM (14 mg/mL) and frozen in liquid nitrogen for long-term storage.



a.



b.

Figure 2. a) Reversed-phase HPLC of *E. coli* D57C CheY phosphonomethylation reaction mixture. The peaks near 30 min are unmodified CheY and the peak at 15 min is phosphono-CheY. b) RP-HPLC of *T. maritima* D54C/C81S CheY phosphonomethylation reaction mixture. The 45 min peak is unmodified CheY and the two peaks at 29 and 31 min are phosphono-CheY.

Phosphonomethylation of *E. coli* D57C CheY

Phosphonomethylation of CheY mutants was carried out using phosphonomethyl triflate (PMT), a compound synthesized as previously reported (*41*). Phosphonomethylation consists of a nucleophilic attack of the cysteine sulfhydryl group (C57 in *E. coli* and C54 in *T. maritima*) on the carbon alpha to the triflate leaving group of PMT (Figure 3).

E. coli D57C CheY stored in 20 mM BES, pH 7.01 was thawed and incubated at room temperature overnight with 10 mM bis-2-mercaptoethylsulfone (BMS). It was buffer exchanged by PD-10 (Amersham) into 750 mM AMPSO, pH 9.01 and the protein concentration measured in matched quartz cuvettes by its absorbance at 280 nm. A molar absorptivity of 8250 M^{-1} cm^{-1}, calculated by the method described in (*46*), was used. Reaction with 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB) in phosphate buffer, pH 7.28 over a 20 min time period was used to estimate the concentration of free thiol in the protein solution (*47*). 950 µL of phosphate buffer and 50 µL of 20 mM DTNB (prepared in the same phosphate buffer) were mixed in a 1 mL plastic cuvette and the absorbance at 410 nm was monitored until a flat baseline was obtained. The protein was added, the solution mixed, and the change in absorbance at 410 nm measured over 20 min. The theoretical ΔA_{410} was approximately 0.100 absorbance units. A ratio of at least 0.75 free thiol concentration to protein concentration was required to proceed with the reaction. The protein was concentrated in a centricon-3 (Millipore) to 1 mM (14 mg/mL). 250 mM $CaCl_2$, $SrCl_2$, or $BaCl_2$ was added from a concentrated stock solution and mixed until homogenous. 120 mM PMT was weighed out on a 5-place balance and dissolved in an amount of 200 proof ethanol equal to half the volume of 336 mM triethylamine (Et_3N). Triethylamine at 336 mM was then combined with the PMT/EtOH mixture and quickly added to the protein, mixing with the pipet tip. The reaction was let proceed at room temperature for 30 – 45 minutes

and was stopped with the addition of 5-10 mM DTT. The precipitate was spun down and the supernatant passed through a PD-10 column for buffer exchange. Protein concentration and thiol concentration were again measured to estimate the extent of modification. RP-HPLC was used to monitor the appearance of the phosphono-CheY peak, which had been previously identified through mass spectrometry and X-ray crystallography (*13*) (Rick Dahlquist and Chris Halkides, unpublished results) as well as phosphorus NMR (Chris Halkides, unpublished results) and total phosphate assay (*48*). Yields between 60% and 90% were desirable and most often achieved. One particular experiment was done differently on *E. coli* CheY to increase the percent alkylation through altering the technique of addition of the PMT/EtOH/Et_3N mixture. S '05 J was prepared as described above. A flea-size stir bar was added to the 1.5 mL vial containing the protein at 10.2 mg/mL and $SrCl_2$ at 200 mM and the solution placed on a stir-plate. 120 mM PMT was dissolved in EtOH and 2.8 equivalents of Et_3N were added. The mixture was quickly pipetted into the stirring protein mixture and reaction proceeded for a half hour. The alkylation was analyzed by DTNB and RP-HPLC as described above.

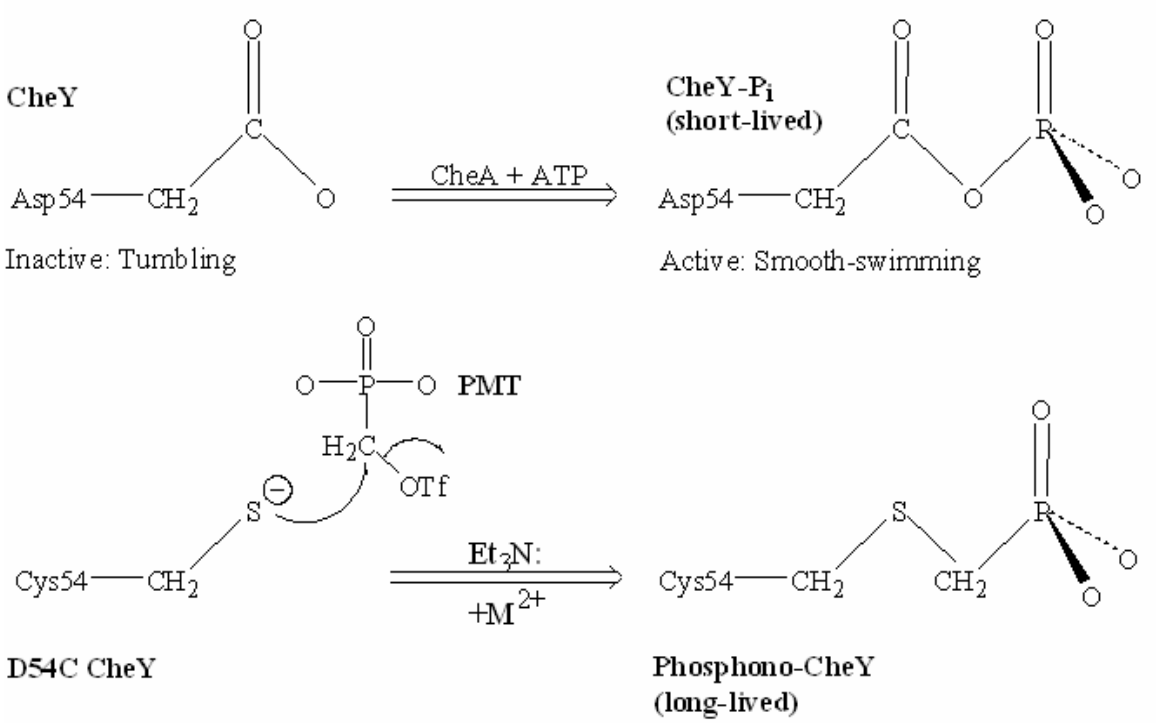


Figure 3. *In vivo* activation of CheY occurs by transfer of a phosphoryl group from ATP to the active site Asp by the action of the histidine kinase CheA, resulting in a mixed anhydride. Phosphonomethylation of *T. maritima* D54C/C81S CheY consists of a nucleophilic attack of the cysteine sulfhydryl group on the methyl carbon of PMT with triflate (OTf) as the leaving group.

Purification of *E. coli* Phosphono-CheY

E. coli phosphono-CheY was purified with a biotin/avidin protocol that yielded approximately 25 – 40% of the amount of protein initially reacted with PMT. After phosphonomethylation, the protein was buffer exchanged into 50 mM BES, pH 7.01 or 50 mM AMPSO, pH 8.50 depending on the biotin reagent used. PEO-Maleimide biotin (Pierce and G-Biosciences) was used in BES and PEO-Iodoacetyl biotin (G-biosciences) was used in AMPSO. 2.5 mM biotin reagent (or 10-fold molar excess over protein) was weighed on a 5-place balance and added to the protein. The reaction was allowed to proceed at room temperature for 12 – 18 hrs. The reaction with PEO-iodoacetyl biotin was done in the dark to minimize reactive iodine radicals. The reaction was terminated with the addition of 10 mM DTT and the protein buffer exchanged into 50 mM BES, pH 7.01. Protein concentration and thiol concentration were measured by absorbance at 280 nm and reaction with DTNB, respectively. A 5 mL immobilized monomeric avidin column (Pierce) was poured under flowing conditions and washed with 5 column volumes of phosphate-buffered saline, pH 7.4. Irreversible binding sites were then blocked with 5 column volumes of 2 mM D-biotin. Excess biotin was rinsed off with 5 column volumes of 0.1 M glycine, pH 2.8 and the column equilibrated in the starting buffer (50 mM BES, pH 7.01 with 1 mM EDTA and 0.02% azide). The protein previously buffer exchanged into the starting buffer was loaded onto the column, with pauses in flow-through every 2.5 mL to maximize binding. 15 mL of effluent was collected directly into a centriprep concentrator (Millipore). Oftentimes, if the amount of biotinylated CheY was greater than 1 mg, the protein was split and put over two avidin columns, again to maximize binding and removal of biotinylated CheY. 15 mL of effluent from each avidin column would be concentrated and then

combined at a smaller volume. Purified phosphono-CheY was then labeled, frozen under liquid nitrogen, and stored at -80 °C.

The cibacron blue column used in the preparation of *E. coli* D57C CheY was tried in the purification of phosphono-CheY. S'05 F' in reducing agent was taken through a PD-10 column into cibacron buffer, pH 7.91. Matched quartz cuvettes were used to determine protein concentration and the mixture was applied directly to a 3 mL Blue column equilibrated in the chloride form. The column was periodically stopped to ensure diffusion of CheY to all binding sites. 10 column volumes of low salt buffer were passed through before the addition of buffer with 2 M NaCl. Absorbance at 280 nm was measured after every 1 column volume collected. The fractions of interest were the first 3 column volumes of low salt buffer and the first 5 column volumes of high salt buffer, and these were collected and concentrated separately. Both the low salt fractions and the high salt fractions were analyzed by RP-HPLC.

Production and Purification of *T. maritima* D54C/C81S CheY

T. maritima D54C/C81S CheY cloned into a pet28a(+) vector (Novagen) encoding an N-terminal hexahistidine-tag was generously provided by the laboratory of Brian Crane at Cornell University. The plasmid was transformed into *E. coli* cell line B834 (DE3) by standard electroporation methods (Appendix A). Selection with 30 µg/mL kanamycin was used to identify colonies that had incorporated the plasmid, and test growths were performed to demonstrate proper IPTG induction before the cell lines were frozen in 2x freezing medium.

Two 5 mL overnight growths in Lauria broth containing 30 µg/mL kanamycin were inoculated with transformed *E. coli* cell line B834 (DE3). The growths were incubated at 37 °C overnight shaking at ≥ 200 RPM. Using sterile technique, two 1 L growths in Terrific Broth

containing 30 µg/mL kanamycin were inoculated with 1 mL from the overnight cultures and incubated at 37 °C shaking at ≥ 200 RPM. The growths were monitored by the absorbance at 600 nm until the optical density reached 1 – 1.5, at which point protein production was induced with 0.2 mM IPTG. The flasks were kept shaking at 37 °C for 4 – 7 hrs. The cells were harvested at 8,000 RPM for 20 minutes at 4 °C in a Beckman JA-10 rotor. The supernatant was poured off and the cell pellets were either frozen on dry ice and kept at -80 °C or they were prepared for sonnication.

The cells were gently suspended in less than 5 mL of Lysis buffer (Appendix B) per 1 gram of wet cell weight. 1.0 mM PMSF and 10 mM 2-ME were added (the Nickel column is not compatible with EDTA, greater than 1 mM DTT, or greater than 10 mM 2-ME). The suspension was sonnicated in a stainless steel beaker until homogenous. Cells were checked for lysis by oil immersion light microscopy at 1000x magnification. During sonnication, cells were chilled to less than 3 °C between bursts and the temperature increase monitored and recorded after each burst. When the cells were broken, the lysate was spun at 18,000 RPM in a JA-20 rotor for 65 min at 4 °C. It was determined by SDS-PAGE that the supernatant contained *T. maritima* D54C/C81S CheY.

A 4 mL (8 mL of 50% slurry) Ni^{2+}-NTA column (Novagen) was poured under flowing conditions and equilibrated in greater than 10 column volumes of lysis buffer. The sonnication supernatant was syringe filtered onto the column and the load collected. The column was then rinsed with 2 – 5 column volumes of wash buffer (Appendix B) and the effluent collected (this usually contained a small amount of CheY and could be re-run over the Nickel column). Bound protein was then eluted with ~ 10 mL of elute buffer (Appendix B). The ideal elution volume was less than 3% of the volume of the size exclusion column used in the final step of

purification. SDS-PAGE was used to ensure the presence of *T. maritima* D54C/C81S CheY in the elute.

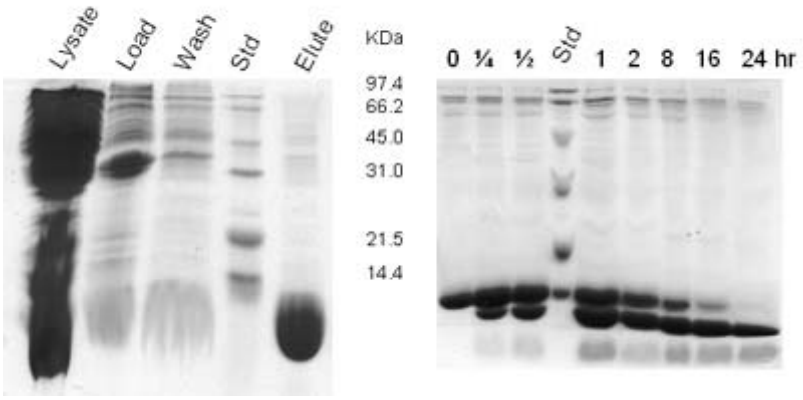
At this point, two methods were used to prepare the protein for thrombin digestion and removal of the his-tag. Initially, the protein was removed from imidazole by dialysis into thrombin cleavage buffer. A 500 mL dialysis was set up and changed at least 2 times. 5 mM DTT was added as a solid (3.9 mg in 500 mL) to each fresh dialysis, although this interferes slightly with thrombin activity. 5-10 mM 2-ME was occasionally used instead. In the later preparations, 5 mM $MgCl_2$, 2.5 mM $CaCl_2$, and 10 mM 2-ME were added directly to the elute from the Nickel column. Thrombin is susceptible to both DTT and 2-ME at concentrations greater than 1 and 10 mM, respectively. Reducing agents disrupt the disulfide bridges that stabilize thrombin's tertiary structure and hence limit its activity.

A Bradford assay or measure of absorbance at 280 nm was used to determine protein concentration. Using restriction grade thrombin from Novagen, 1:8,000 w:w (thrombin**:**protein) or 0.25 units per mg of protein was added. Using high purity grade thrombin from MP Biomedicals, 1:3000 w:w (thrombin:protein) or 0.5 activity units (NIH units) per mg of protein was added. Cleavage trials to determine the best ratio thrombin:protein were performed any time a new type of thrombin was used. The reaction proceeded overnight (at least 16 h) rotating at room temperature. SDS-PAGE was used to check if digestion was complete. A small sample taken before digestion provided a band with which to compare the digested product. 1 mM PMSF was added to inhibit further thrombin activity after digestion was complete.

Size-exclusion chromatography was used to separate CheY from thrombin and the his-tag. The Cornell protocol from which this was adapted employs a Superdex 200 column (GE Healthcare) run under fast protein liquid chromatography conditions, a technique analogous to

high performance liquid chromatography but with much larger columns for large scale purification. In order to best mimic their separation, Sephadex G-50 was selected. The MW cutoff is slightly below the MW of thrombin and therefore it elutes in the void volume, and sufficient separation between pure CheY and his-tag (2 KDa) is achieved. A 466 mL column 95 cm high and 2.5 cm in diameter was poured under flowing conditions as uniformly as possible at or below the temperature at which it was run. Scopes (*49*) and Hirs (*50*) offer tips on pouring gel filtration/SEC columns. The column was equilibrated with 4 column volumes of buffer, pH 7.0 – 7.5. MOPS was used initially, although due to its high cost, a switch was made to phosphate buffer at pH 7.2. EDTA at 1 mM and sodium azide at 0.02% w:v were included in the buffers from this point forward. 10 mM 2-ME and 1 mM PMSF were added to the last two liters of buffer used to equilibrate the column as well as the liter of buffer used to elute the protein. For best separation, the volume of the sample to be applied was kept to less than 3% of the column volume, and the sample was loaded in a manner that did not disturb the top of the column bed. A safety loop was set up connecting the inlet of the column to a 1 L reservoir of buffer (same buffer used to equilibrate the column in), and a Pharmacia RediFrac fraction collector was connected to the column outlet. The best resolution was achieved when the column was run at ~ 1/2 mL/min and fractions at 10-12 min intervals were collected in clean glass test tubes. The fractions were collected overnight and analyzed by A_{280} in the morning. Three peaks eluting at approximately 150 mL, 250 mL, and 500 mL of buffer were identified as thrombin, CheY, and his-tag, respectively (Figure 5). The fractions containing CheY were pooled based on A_{280}, SDS-Page, and absorbance spectra. 50 mM 2-ME was added to the fractions believed to contain CheY while awaiting results from the gel. The protein was then concentrated in an Amicon stirred-cell concentrator using a YM3 or PLBC 3,000 MW cutoff membrane (Millipore),

aliquoted, and frozen in liquid nitrogen. Occasionally, the protein was dialyzed into 50 mM Tris-HCl, pH 7.5 for storage. All aliquots were labeled and stored in the -80 °C freezer.



a. b.

Figure 4. a.) SDS-PAGE of the purification of 6xHistidine-tagged *Thermotoga maritima* D54C/C81S CheY by immobilized metal affinity chromatography. Ni^{2+}-NTA has affinity for histidine residues and is used for the first purification step of CheY from the soluble cell lysate. The lysate in lane 1 contains all of the soluble protein found after sonnication of the cells. Lane 2 is the protein that did not bind to the nickel column upon loading. Lane 3 is the protein that bound initially but was washed off easily with low concentrations of imidazole. Lane 5 is the protein that eluted under increased imidazole concentration (200 mM). Imidazole competes with histidine-tagged CheY for the nickel binding sites and therefore CheY elutes under high imidzaole concentration. b.) SDS-PAGE of the cleavage reaction used to remove the 6xHistidine tag from the N-terminus of *T. maritima* D54C/C81S CheY. A linker region containing nucleic acids that encode for the amino acids LVPRGS was engineered into the plasmid that codes for the *T. maritima* D54C/C81S CheY protein. Thrombin cleaves between R and G and separates the Histidine tag from CheY, leaving the amino acids GSH present on the N-terminus of the protein. This cleavage reaction can be followed by SDS-PAGE.

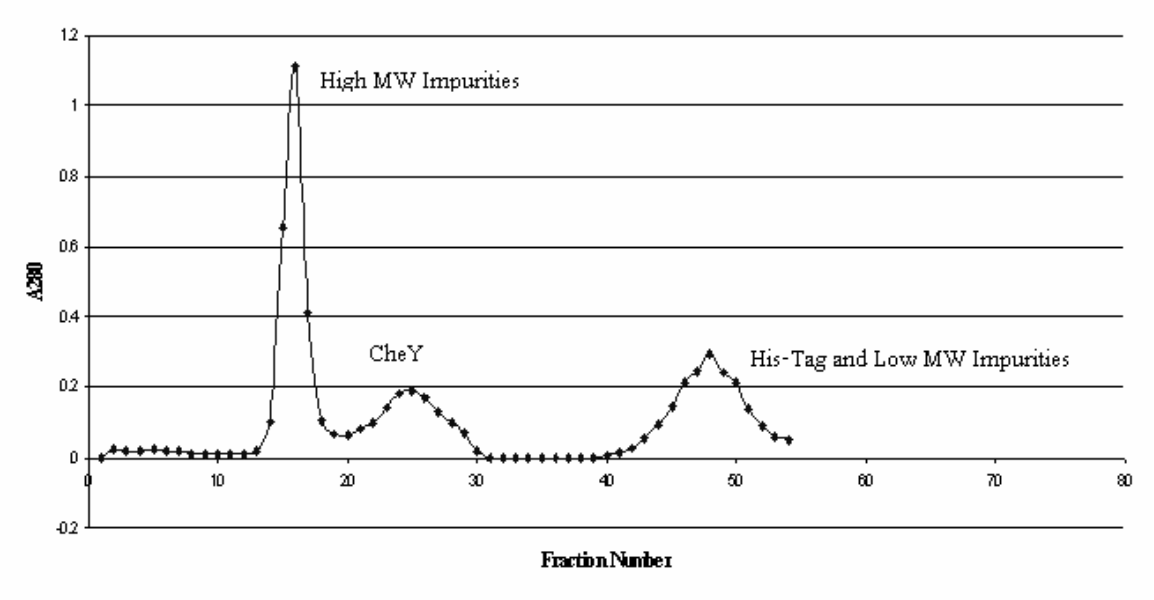


Figure 5. Size-exclusion chromatography is used as the final step in purification of *T. maritima* D54C/C81S CheY. Thrombin and higher molecular weight impurities (>30 KDa) elute in the void volume (~150 mL), CheY elutes at ~250 mL, and the his-tag elutes near 500 mL. The identity of each peak was confirmed by SDS-PAGE (see Figure 6).

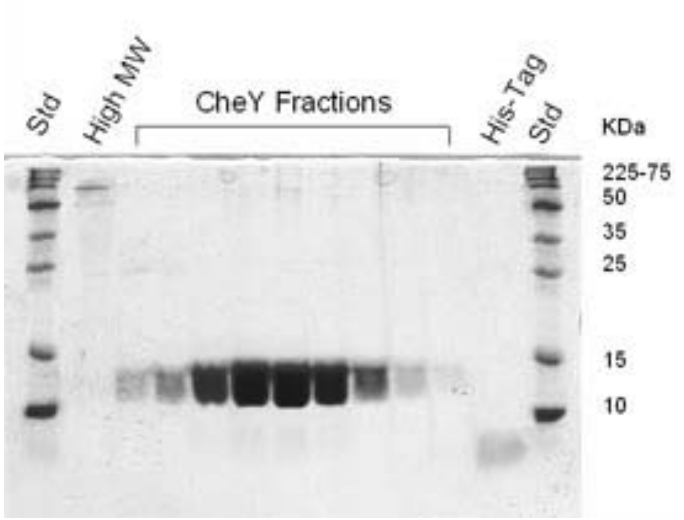


Figure 6. SDS-PAGE of selected fractions of the eluate from the size-exclusion column shown in Figure 5. Lane 2 is the protein that elutes in the first peak of Figure 5. Lanes 3-11 are the fractions representing the second peak. The protein has a molecular weight consistent with *T. maritima* D54C/C81S CheY (13.5 KDa). Lane 12 is the protein that elutes in the third peak in Figure 5. Experiments were able to identify this as the His-tag.

Phosphonomethylation of *T. maritima* D54C/C81S CheY

T. maritima D54C/C81S CheY stored in 50 mM phosphate buffer, pH 7.2 was thawed and incubated at room temperature overnight with 10 mM bis-2-mercaptoethylsulfone (BMS). It was buffer exchanged by PD-10 (Amersham) or two Penefsky spin columns into 250 mM AMPSO, pH 9.01 or 250 mM TAPS, pH 8.50 and the protein concentration measured in matched quartz cuvettes by its absorbance at 280 nm. A molar absorptivity of 2560 M^{-1} cm^{-1}, calculated by the method described in (*46*), was used. When spin columns were employed for buffer exchange, Bradford assay was the method of choice for protein quantitation. Reaction with DTNB over a 40 min time period was used to estimate the concentration of free thiol in the protein solution. A ratio of 0.75 free thiol concentration to protein concentration was desired in order to proceed with the reaction. The protein was concentrated in a centricon-3 concentrator (Millipore) to between 6 and 10 mg/mL. Divalent and trivalent metals at final concentrations ranging from 1 mM to 250 mM were added from a concentrated stock solution and mixed until homogenous. 120 mM PMT was weighed out on a 5-place balance and dissolved in an amount of 200 proof ethanol equal to half the volume of 360 mM triethylamine. Three equivalents of triethylamine (360 mM if PMT is 120 mM) was then combined with the PMT/EtOH mixture and quickly added to the protein, mixing with the pipet tip. The reaction was let proceed at room temperature for 30 – 45 minutes and was stopped with the addition of 5-10 mM DTT. The precipitate was spun down and the supernatant passed through a PD-10 column (or two Penefsky columns) for buffer exchange. Protein concentration and thiol concentration were again measured as described above to estimate the extent of modification. It is important to note that phosphonomethylation reactions were carried out using a range of pH, buffer concentrations, metals, metal concentrations, and PMT concentrations. Triethanolamine and KOH were

occasionally substituted for triethylamine. Buffer exchanges and analyses were all performed as described above. RP-HPLC samples of phosphonomethylation reaction mixtures are referred to as single-prime samples (ex: A').

Purification of *T. maritima* Phosphono-CheY by Biotinylation/Avidin

T. maritima phosphono-CheY was purified by the biotin/avidin protocol described above using both PEO-Iodoacetyl biotin and PEO-Maleimide biotin. A third type of biotin, Biotin-HPDP (G-Biosciences) was also used (*51*). Time points of biotinylation reactions with unmodified CheY were taken to describe their kinetics before the method was used to purify phosphono-CheY. The reaction was monitored both by DTNB for loss of free thiol and by RP-HPLC for the disappearance of unmodified CheY and the appearance of biotinylated CheY. PEO-iodoacetyl biotin and PEO-maleimide biotin were used under the standard conditions described above for *E. coli* phosphono-CheY. Biotin-HPDP, which is insoluble in water, was prepared as a 50 mM stock solution in DMSO with stirring and minimal heating in a small glass vial. The protein solution was buffer exchanged into 20 mM Tris, pH 8.02 with 1 mM EDTA and 0.02% azide. A 20-fold dilution upon addition of Biotin-HPDP to the protein solution was used, affording a 2.5 mM final concentration of biotin and 5% v:v DMSO. The reaction was followed as close to theoretical completion by absorbance at 343 nm using a molar absorptivity of 8080 M^{-1} cm^{-1} for the resulting pyridine-2-thione leaving group, and the data were analyzed in detail as described in the results. The mixture was also analyzed by RP-HPLC for the disappearance of unmodified CheY and the appearance of biotinylated CheY before passing over avidin. RP-HPLC samples of biotinylation reaction mixtures are referred to as double-prime samples (ex: A'').

Two types of immobilized avidin were used to remove biotinylated *Thermotoga* CheY. The procedure for the monomeric avidin column is the same as described for *E. coli* phosphono-CheY. A 5 mL column of tetrameric avidin resin (immobilized avidin, Pierce) was poured under flowing conditions and immediately equilibrated in 10 column volumes of 50 mM BES, pH 7.01 with 1 mM EDTA, 0.02% azide, 1 µM pepstatin A and 1 µM leupeptin. The protein, buffer exchanged into the same buffer, was loaded onto the column. 1.5 mL was passed into the column bed and incubated for 15 min. The rest of the protein was then passed into the column bed, again incubating for 15 min. 250 µL of 50 mM BES, pH 7.01 with 1 mM EDTA, 0.02% azide, 1 µM pepstatin A and 1 µM leupeptin was added to force the remaining protein into the column bed. After 1 hour total incubation, 18 mL of eluate was collected continuously directly into a centriprep concentrator. The sample was the concentrated to greater than 1.5 mg/mL for analysis by RP-HPLC. RP-HPLC samples of purified phosphono-CheY samples are referred to as triple-prime samples (ex: A’’’).

Purification of *T. maritima* Phosphono-CheY by HPLC

Three types of high performance liquid chromatography were employed in attempt to purify *T. maritima* phosphono-CheY: reversed phase-HPLC (RP-HPLC), hydrophobic interaction chromatography-HPLC (HIC-HPLC), and anion-exchange chromatography. In each case, peaks were collected downstream of the variable wavelength detector. Protein purified from reversed phase was quick-frozen in a pear-shaped flask and placed in the lyophilizer overnight at -50 °C and between 100-700 millibars pressure. When all solvent had evaporated, the flask was removed. The powder was then re-suspended in the desired buffer or solvent. Purification of phosphono-CheY by HIC-HPLC first required biotinylation of unmodified CheY

in phosphonomethylation reaction mixture. Biotinylation was performed with PEO-iodoacetyl Biotin according to the standard procedure. The protein mixture was then concentrated to a volume less than 300 µL and combined with 200 µL of saturated (3.8 M) ammonium sulfate. The mixture was injected and protein eluted on a 0 – 90% (1.6 M to 0.16 M ammonium sulfate) gradient over 30 minutes. One minute fractions were collected and buffer exchanged into 50 mM BES, pH 7.01. Phosphono-CheY was purified from fractions 14-17.

Purification of *T. maritima* Phosphono-CheY with Immobilized Glutathione

Before phosphono-CheY was tested, a sample of unmodified CheY was buffer exchanged into 50 mM phosphate buffer, pH 7.28 with 1 mM EDTA and 0.02% azide. 855 µg of protein was allowed to react with a 20-fold molar excess of DTNB for approximately 2 hours. Absorbance at 410 nm using a molar absorptivity of 14150 M^{-1} cm^{-1} for the TNB anion was used to monitor completeness of the reaction. The mixture was buffer exchanged into 20 mM Tris, pH 8.02 with 1 mM EDTA and 0.02% azide. A 10 mL immobilized glutathione (Pierce) column in a reduced state was poured from slurry under flowing conditions and washed with at least 10 column volumes of 20 mM Tris, pH 8.02 with 1 mM EDTA and 0.02% azide until no DTNB reactivity of the effluent was observed. The protein mixture was loaded onto the column until the meniscus of the solution was touching the top of the column bed. The column was then rotated at ~100 RPM for 24 hours at room temperature to allow glutathione to react with all of the TNB-labeled CheY. The column was washed with three column volumes (30 mL) of 20 mM Tris, pH 8.02 with 1 mM EDTA and 0.02% azide and the effluent collected. The CheY attached to glutathione was then eluted with 3 column volumes of Tris buffer with 20 mM 2-ME and the

effluent collected. Both the wash and elute were concentrated and analyzed by Bradford assay, RP-HPLC and SDS-PAGE for the presence CheY.

In-Gel Digestion

Protocols for in-gel digests were obtained from the laboratory of Dan Liebler at Vanderbilt University. Handling and operation of pre-made gels were done according to manufacturers recommendations. Pre-cast NuPage Bis-Tris gels (10% and 12%), MOPS SDS Buffer Kit, and Colloidal Blue staining kit were all purchased from Invitrogen. The Bis-Tris gel package includes 4x LDS sample buffer, 500 mM DTT, and an antioxidant. The Mighty Small II SE250 gel running box was purchased from Hoefer. *OneQuant* Iodoacetamide was purchased from G-Biosciences and prepared immediately before use according to manufacturer's recommendations. Sequencing grade modified trypsin, frozen was purchased from Promega. Molecular weight standards were purchased from BioRad (#161-0304) and EMD Biosciences (# 69079). BioRad Low Molecular Weight Standards were prepared with 5 µL standard, 25 µL 4x sample loading buffer, and 70 µL DI water and were frozen. EMD Biosciences Perfect Protein Markers came ready to use and were frozen in 10 µL aliquots. NuPage Bis-Tris gels, LDS sample buffer, NuPage 10x Reducing Agent (0.5 M DTT, 98% Research Grade), and antioxidant are stored at 4 °C and removed immediately prior to use. MOPS SDS Running Buffer was stored at room temperature as a 20x concentrate.

The Bis-Tris gel cassette was removed from its pouch and rinsed with DI water. The white tape at the bottom of the gel cassette was peeled off and the comb at the top of the gel cassette was carefully removed. The wells were then rinsed 2-3x with 1x running buffer, inverting and shaking each time to remove the buffer. The gel was oriented in the gel running

box with the well-side of the cassette facing inwards, exposed to the upper buffer reservoir. The gel was clamped in place and the tightness of seal was checked by adding 1x running buffer to the upper reservoir first and checking for leaks. The upper reservoir holds approximately 60 mL of running buffer. 150 µL of NuPage antioxidant was added to this reservoir prior to running.

The sample to be digested was prepared by taking approximately 20 µg of protein for each desired gel lane, and combining it with either DI water or phosphate buffered saline, pH 7.4 and 4x LDS sample buffer (to a 1x final concentration). If purified by reversed-phase conditions, the protein was quantified either by colorimetric assay or by absorbance at 280 nm. The protein was then aliquoted so that upon lyophilization a known amount of sample was dissolved and loaded into a gel for a more accurate protease:protein ratio. Samples were reduced with a 10x dilution of NuPage Reducing Agent and heated at 70 °C for 10 min. Samples of *Thermotoga maritima* D54C/C81S CheY were heated to 100 °C for 10 min due to their heat-stability. 20 µL was loaded into each well, along with 2 molecular weight standards on either end. The gels were run at a constant 200 V for approximately 40 min. The current started at ~110 mA and finished at ~70 mA.

When the run was complete, the gel cassette was removed by physically separating the two plates, and it was placed into a container appropriate for staining. The gel was rocked at 1 RPM in fixing solution (40 mL DI water, 50 mL methanol, and 10 mL acetic acid) for 10 min and the solution discarded. The gel was then rocked at 1 RPM in staining solution (55 mL DI water, 20 mL methanol, and 20 mL NOVEX Colloidal Blue Stainer A) for 10 min. 5 mL of Stainer B was added and the gel rocked for 4-5 h. The staining solution was removed, and the gel washed once with water for 45 min and then again in water overnight to destain.

The desired protein bands were excised from the gel and cut into 1 mm cubes. Each band was placed into a 1.5 mL microcentrifuge tube (bands can be combined at this point), 100 µL of 100 mM ammonium bicarbonate (Ambic) was added and the bands incubated for 15 min. The liquid was discarded and 150 µL of 100 mM Ambic and 1.67 µL of 300 mM DTT were added and the samples incubated at 50 °C for 15 min. The samples were then removed from heat and 2 µL of 1 M iodoacetamide added, and the protein was incubated in the dark at room temperature for 15 min. The liquid was discarded and 100 µL of 50:50 acetonitrile:50 mM Ambic (v:v) added. This was repeated at least three times until the blue hue was gone. 50 µL of 100% acetonitrile was added, and the samples were let sit until the slices turned white. Liquid was discarded and the samples were dried in a vacuum centrifuge for 10 min.

The dehydrated gel slivers were re-suspended in 50 µL 25 mM Ambic and 2 µL of 0.2 mg/mL trypsin (1:50 w:w protease:protein) and incubated at 37 °C overnight (~24 h). The liquid was pipetted into another microcentrifuge tube and set aside. 50 µL of 1% formic acid in 60% acetonitrile was added to each sample and incubated at room temperature for 15 min. The extract was removed and added to the vial set aside. This was repeated several times and the combined liquid was evaporated using vacuum centrifuge. The peptides were dissolved in 1% formic acid and cleaned purified over a C-18 reversed-phase minicolumn prepared in a micropipette tip (Millipore and Varian) according to protocol (Appendix C). Samples can be directly injected on LC/MS or dried in a vacuum centrifuge and dissolved in an appropriate solvent.

Mass Spectrometry

Mass spectrometry was done at the Center for Marine Science using electrospray ionization and time-of-flight mass analysis on a Q-Star XL MS/MS (Applied Biosystems) connected to an HP1100 series binary pump with autosampler. Mobile phase A was 98% water, 2% ACN, 0.1% formic acid (FA), and 0.01% TFA and mobile phase B was 2% water, 98% ACN, 0.1% formic acid (FA), and 0.01% TFA. The flow-rate of the system was 200 µL/min and injection volumes were 3 µL. Samples for injection were taken to dryness to remove any excess TFA from our normal RP-HPLC conditions and prepared for injection by re-suspension in a small volume of mobile phase A. Standard direct injections were made using 100% mobile phase B. Gradient LC-MS employed a Vydac 2.1 mmID x 250 mm C-18 column (described above) with a 5 minute isocratic elution of 100% mobile phase A, followed by a 30 minute gradient of 100% A to 100% B. A 30 minute equilibration at 100% A was necessary between runs. Blanks consisting of mobile phase A were run prior to any samples as well as between samples.

Mass spectrometry was performed on intact protein as well as tryptic and LysC digestions. Tryptic digests were performed in-gel and prepared for MS as described. Bovine serum albumin was digested first for practice, and *T. maritima* D54C/C81S CheY, phosphono-CheY, and a 25 KDa impurity present in Cornell CheY were also digested. Intact protein was prepared for ESI-TOF mass spectrometry analysis by purification from RP-HPLC followed by lyophilization. Lyophilizate of the CheY*, CheY, and phosphono-CheY peaks, as well as the uneven baseline before phosphono-CheY (see Figure 40) were each re-dissolved in 98% water, 2% ACN, 0.1% formic acid (FA), and 0.01% TFA and analyzed by ESI-TOF MS.

LysC digestion mixtures were prepared according to the methods outlined in Peptide Mapping and analyzed in a variety of ways. LysC digestions of Cornell CheY and UNCW CheY (see methods for Peptide Mapping) were injected in triplicate onto the 4.6 x 250 mm C-18 column for collection using the standard ACN/TFA mobile phases. A baseline sample was collected for a negative control, an arbitrary peak was collected for a positive control, and the peak of interest was collected for identification. These six samples (three from Cornell CheY and three from UNCW CheY) were dried in a vacuum centrifuge to ~20 uL and analyzed by direct injection ESI-TOF mass spec.

LysC digestion of the two peaks from pool T (see methods for Peptide Mapping) were prepared in 5% ACN/0.1% TFA and injected onto a 10 mm ID C-18 column to clean up the digest and remove urea and any buffer salts. Upon injection mobile phase A (5% ACN, 0.1% TFA in water) was allowed to flow for ~12 minutes or until the baseline stabilized. Flow was switched to 100% mobile phase B (80% ACN, 0.8% TFA in water) and the effluent collected. The effluent was then frozen over dry ice/acetone, lyophilized to powder, and submitted for ESI-TOF MS analysis. The digestion mixtures of the two peaks from M pool 2 were not cleaned up by RP-HPLC and submitted as is (see methods for Peptide Mapping). LysC digestion and MS preparation of phosphono-CheY was done exactly as described for the digestions of M pool 2.

Mass spectrometry was also performed in collaboration with labs at Cornell, UC Santa Barbara, and Duke University. MALD/I-MS performed at Cornell was mostly aimed at characterizing CheY* but was also done to confirm the phosphonomethyl modification of *T. maritima* D54C/C81S CheY. ESI-TOF performed at UCSB was used exclusively to characterize CheY*. MALD/I-MS and LC/ion trap MS at Duke was used to identify CheY peptides and confirm the phosphonomethyl modification of *T. maritima* CheY.

Samples sent to Cornell were prepared for MALD/I-MS analysis. The impurity peak (CheY*) and the CheY peak present in R pool 1 and O pool 1 were collected from RP-HPLC and evaporated down to ~20 pmol/µL in a vacuum centrifuge. The same was done with CheY peak from Cornell CheY. A sample of purified putative phosphono-CheY as well as a portion of O pool 1 was buffer exchanged through 2 Penefsky columns into 10 mM ammonium acetate. All samples were frozen under liquid nitrogen for shipment on dry ice.

Samples sent to UCSB were prepared for ESI-TOF MS analysis. First, two pools of *T. maritima* D54C/C81S CheY prepared in-house and possessing different RP-HPLC peak area ratios of CheY*:CheY were removed from -80 °C and split in half. R pool 1 at 3.5 mg/mL was split into two 750 µL-portions, one of which was dialyzed into deionized water and the other dialyzed into deionized water with 10 mM DTT. The same was done with O pool 1 at 3.9 mg/mL split into 625 µL-portions. Each dialysis was done into 500 mL and one change was made, providing 1333-fold and 1600-fold dilutions of the buffer for R and O, respectively. The samples were removed from dialysis and pipetted into a 1.5 mL microcentrifuge tube and quick-frozen under liquid nitrogen for shipment on dry ice.

A number of samples prepared in a variety of ways were sent to Duke University. Tryptic digestions of CheY and phosphono-CheY were prepared as described in the methods for in-gel digestion. LysC digestions of phosphono-CheY and the two peaks from N pool 2 were prepared as described above. The 20 min doublet was analyzed in-tact and was prepared as described above. Peaks collected in ~1.6 M ammonium sulfate from HIC-HPLC were buffer exchanged by PD-10 into 20 mM ammonium bicarbonate and concentrated in centricon-3 concentrators down to ~2 mg/mL. The samples were quick-frozen under liquid nitrogen for shipment.

Characterization of CheY*

Peptide Mapping

Two peptide mapping experiments of *T. maritima* D54C/C81S CheY were performed. In the first, a volume containing 4 mg of CheY was removed from R Pool 1, a sample that showed two peaks by RP-HPLC (see Figure 14c) and from Cornell CheY, a sample that showed only one peak by RP-HPLC. Each was diluted 10-fold with 8 M urea, reduced with 10 mM DTT and concentrated in centricon-30 concentrators to a final protein concentration of 5 mg/ml. The protein was frozen in a dry-ice/acetone bath for digestion the next day. Protein was thawed and added to LysC buffer (4:1 8 M Urea:500 mM Tris) to a final volume of 2.9 mL and concentrated down to 3 mg/mL in centricon-30 concentrators. Protein quantitation was done before concentration by absorbance at 280 nm. Protein was concentrated to 3 mg/mL in a volume of ~300 uL and lysyl endopeptidase (Wako Pure Chemical Industries) was added 1:50 w:w protease:protein. Nitrogen gas was blown over the samples and they were incubated for 16.5 hrs at 37 °C. The digests were then centrifuged 10K RPM for 10 min, flushed with $N_{2\ (g)}$ and placed on wet ice for sampling. 60 µg were removed from each digest and prepared immediately in 5% ACN/0.1% TFA for RP-HPLC analysis. A control experiment consisting of the protease with no protein was taken through the entire process alongside the two digestions. A gradient of 0% MPB to 100% MPB was used to map the resulting LysC peptides. Mobile phase A was 5% ACN/0.1% TFA in MilliQ, and mobile phase B was 80% ACN/0.08% TFA in MilliQ water. This experiment was repeated later with R pool 1, M pool 2, and Prep D.

Another peptide mapping experiment was performed on protein purified from RP-HPLC. In two separate experiments, pool T and N pool 2 were injected onto the 4.6 mm ID C-18 column and the CheY* peak and the CheY peak were collected individually and lyophilized to

powder. Lyophilizate was re-suspended in 8 M Urea and taken through the procedure exactly as described above.

Reductions

O pool 1, which yielded two peaks by RP-HPLC, at 3.9 mg/mL in 50 mM MOPS, pH 7.2 was separated into five 250 µL aliquots. Each was let stand open for 25 minutes and sampled for RP-HPLC and DTNB analysis. Four of the aliquots were reduced with a different reducing agent and the fifth aliquot was left unreduced and capped for a control. 5 mM TCEP, 5 mM BMS, 5 mM DTT, and 22 mM 2-ME were used as reducing agents. Each reduction was sampled at 1, 2, and 20, and 26 hours for RP-HPLC analysis by preparing each sample in 35% ACN and immediately freezing in a dry-ice/acetone bath. Analysis was done under the standard shallow gradient for *T. maritima* CheY (46.5% to 54% B over 50 min).

Counter-ion Effect

O pool 1 was prepared in three different buffers with 2 different concentrations of TFA in order to test whether the counter-ion in RP-HPLC conditions was leading to multiple peaks in pre-phosphonomethylated CheY samples. HPLC samples were prepared with protein taken from 50 mM MOPS, pH 7.2, 50 mM Tris-HCl, pH 7.5 with 150 mM NaCl, and 50 mM phosphate buffer, pH 7.2 with 1 mM EDTA. Samples were prepared to final concentrations of 0.1% TFA and 0.4% TFA in 35% ACN and injected on the 4.6 x 250 mm C-18 column under the standard shallow gradient for *T. maritima* CheY.

His-tagged Protein

Two experiments were aimed at determining whether CheY* was residual His-tagged CheY. First, a thrombin digestion of His-tagged CheY from the Ni-NTA eluate was followed by RP-HPLC. The reaction was sampled before addition and at 3, 5, 9 and 23 hours. Samples were

immediately prepared in 35% ACN and frozen. They were thawed immediately prior to injection. Second, O pool 1, or protein that had supposedly already had the his-tag removed, was exchanged by PD-10 into lysis buffer and passed back through a Ni-NTA column according to the standard protocol outlined in the preparation of *T. maritima* D54C/C81S CheY. The load, wash, and elute were monitored by RP-HPLC for the presence of CheY and absence CheY*.

Separation by Anion Exchange

A 5 mL DEAE cellulose column was poured at 3 °C under flowing conditions and equilibrated in its chloride form with 10 column volumes of 2 M NaCl in 50 mM Tris, pH 8.07. Tris has a $\Delta pK_a/\Delta T$ = -0.028 so Tris pH 8.07 at 22 °C is pH 8.6 at 3 °C. The column was subsequently rinsed with 50 mM Tris, pH 8.07 until the conductivity dropped to its initial value of 3.30 mS. O pool 2 at 1.7 mg/mL in 2.5 mL was loaded onto the column at a flow rate of 1/3 mL/min and the load collected. 50 mM Tris pH 8.07 was passed through the column and 1.5 mL fractions were collected. Fractions 2 and 3 were combined and concentrated and analyzed by A_{280} for the presence of protein. The column was then washed with 2 M NaCl in 50 mM Tris, pH 8.07 and 1.5 mL fractions were collected. The fractions were scanned for protein the first four combined and concentrated. Both the low salt and high salt fractions were analyzed separately by RP-HPLC.

Quantitation of *T. maritima* D54C/C81S CheY

There are a number of reported methods for calculating molar absorptivites of proteins at 205 and 280 nm. Scopes outlines one briefly in “Protein Purification” and Deutscher the same in “Guide to Protein Purification.” Scopes outlines the method of calculating the molar absorptivity of a protein at 205 nm (*52*), and an informative method for estimating the molar

absorptivity of a protein at 280 nm is described by Iersel, et al. (*53*). Both reports derive equations that enable manipulation of measured absorbance values at 205 nm and 280 nm to determine a protein's molar absorptivity.

The method was first attempted on standard solutions of BSA and egg albumin and then on a sample of *T. maritima* CheY. The measurements were repeated three times on two different spectrophotometers, a single beam Pharmacia Ultraspec 2000 and a double beam Cary 1E UV/Vis Spectrophotometer (Varian). A 1.0 mg/mL standard solution of BSA was prepared in a 10 mM Tris, pH 7.48, 0.1 mM EDTA, in MilliQ water. Wavelength scans of the protein at 500 µg/mL concentration were taken from 250 nm to 400 nm and the absorbance at 280 nm recorded. It is important that the absorbance at 280 nm be greater than 0.200 but less than 1.00 for the best accuracy. Absorbance at 205 nm is more sensitive and required special cuvettes that did not absorb in this region of the spectrum. Before each scan, the cuvettes were cleaned in glacial acetic acid and rinsed profusely with deionized water. In addition, buffers were kept to a minimum concentration and azide was left out to minimize buffer contributions to the absorbance at 250 nm. First, both cuvettes were filled with 1.0 mL buffer and scanned from 190 nm to 400 nm. This was done three times and it was found that the A_{205} of the sample cuvette is 11 mAU below the reference cuvette. This was only necessary on the Pharmacia Spectrophotometer, since it is a single-beam instrument. A 10 uL aliquot of BSA standard is then added to the sample cuvette and mixed by inversion. Again scans from 190 to 400 nm are performed. The same procedure was done for egg albumin and for *T. maritima* phosphono-CheY sample TMY 2'.

From absorbance values at 205 and 280 nm, the molar absorptivities at each wavelength can be calculated using the following formulae:

$\varepsilon_{205} = 27 + 120 * (A_{280}/A_{205})$ Ref. (*46*)

$A_{280}^{0.1\%} = (27 \times A_{280}/A_{205}) / (1 - 3.85 \times A_{280}/A_{205})$ Ref. (*53*)

ε_{280} = molecular weight * $A_{280}^{0.1\%}$

$A_{280}^{0.1\%}$ is the absorbance of a 1 mg/mL protein solution at 280 nm. Better Bradford and BCA assay (Pierce) were also performed in tandem to determine their accuracy and consistency with the results of the A_{205}/A_{280} method. For the Bradford Assay, a standard curve of absorbances at 595 nm was prepared using various concentrations of egg albumin in 50 mM TAPS, pH 8.5 according to the serial dilution method outlined in the Pierce Better Bradford literature. 1.5 mL of working Better Bradford reagent (Pierce) was added to 50 µL portions of protein and a linear regression of the first four data points was obtained. A 10 µL portion of *T. maritima* phosphono-CheY sample TMY 2' was added to 90 µL of 50 mM TAPS pH 8.5 and 2.0 mL Better Bradford reagent and was analyzed at 595 nm. For the BCA assay, a working BCA reagent was prepared with 1.0 mL of 4% $CuSO_4$ and 50 mL of BCA Reagent (Pierce). A standard curve was prepared using 0 – 51µg/mL of egg albumin in 50 mM TAPS, pH 8.5 with 2.0 mL of working BCA reagent (final volume 2.1 mL). The samples were incubated at 60 °C for 30 min, and linear regression of the first five data points was used to make the standard curve. 10 µL of sample TMY 2' were added to 90 µL of 50 mM TAPS, pH 8.5 and 2.0 mL working BCA reagent, incubated in a heat block at 60°C for 30 min, and absorbance at 560 nm measured.

RESULTS

Production and Purification of *E. coli* D57C CheY

As expected from previous published results (*41*) the majority of D57C CheY was found in inclusion bodies after growth and sonnication. The first urea wash contained strong CheY bands, 2 strong impurity bands at molecular weights 30 KDa and 75 KDa, and a few other faint bands. The second urea wash contained CheY (along with a few faint impurities) on one occasion and was included in the purification. The faint impurities were removed with the pH 7.0 anion exchange column, but the stronger impurities eluted immediately prior to CheY. In one preparation, fractions 15-17 from the column contained the impurities in equal proportion to CheY by SDS-PAGE, but fractions 18-22 contained only slight impurity bands. All fractions that contained CheY were dialyzed and passed over the pH 7.9 cibacron blue column, which was effective in removing the stronger impurities. The two preparations produced 50 and 55 mg of D57C CheY per liter of growth from the insoluble fraction. One preparation produced 15 mg of D57C CheY of equal purity from the soluble fraction.

Phosphonomethylation of *E. coli* D57C CheY

Phosphonomethylation of *E. coli* D57C CheY was carried out on twenty three reactions containing ~185 mg of *E. coli* D57C CheY and resulted in ~37 mg of purified phosphono-CheY. Alkylation percentages varied from 33% to 76% by DTNB and from 36% (Figure 7) to greater than 90% (Figures 6a and 6b) by RP-HPLC peak areas. Percentage yield calculated from DTNB and RP-HPLC agreed within 20% of each other for roughly half of the reactions (Table 1). Only twice was the percent modification by DTNB greater than the percent areas by RP-HPLC (S ’05 D’ and F ’04 A’). The most successful alkylations were done in 750 mM AMPSO, pH 9.01 with

1 mM EDTA and 0.02% azide in the presence of 200 and 250 mM $SrCl_2$, 120 mM PMT, and 2.8 equivalents of triethylamine. PMT 2005.C and PMT 2005.2B worked equally well in these reactions, as did $CaCl_2$ in place of $SrCl_2$.

Slightly less than half of protein was lost during the phosphonomethylation reaction, however, only a quarter of the protein was lost when the buffer concentration was increased to 1.5 M (S '05 C'). The experimental addition of the PMT/EtOH/Et_3N mixture (S '05 J') to a stirring protein/metal mixture also reduced the amount of protein lost. A slight positive correlation was found between precipitation and protein loss, and precipitation was negatively correlated with alkylation percentage. For example, no precipitation was seen in the reaction S '05 J', but the percent yield was only ~40%.

Purification of *E. coli* Phosphono-CheY by Biotinylation/Avidin

E. coli D57C CheY was purified by reacting the unmodified thiol group with biotin and removing the biotinylated CheY through the affinity biotin has for avidin. Biotinylation was initially done with PEO-maleimide biotin in BES at pH 7.0 and PEO-iodoacetyl biotin in AMPSO at pH 9.0. Precipitation along with significant loss of protein often accompanied biotinylation with the maleimide reagent, so the iodoacetyl reagent was used almost exclusively. In order to increase the specificity of biotinylation, the reaction with PEO-iodoacetyl biotin was re-examined at a slightly lower pH (pH 8.5 compared to pH 9.0). A decrease in the free thiol concentration measured by DTNB assay and the disappearance of the D57C CheY peak by RP-HPLC provided a way to measure completeness of the reaction between *E. coli* D57C CheY and PEO-iodoacetyl biotin at pH 8.5. The reaction was analyzed over an 8.5 hour period and was essentially complete by 2.75 hours by both DTNB and RP-HPLC (Table 2). The CheY peak had

virtually disappeared 25 minutes into the reaction according to RP-HPLC (Figure 9), and DTNB measurement indicated only 10% of the initial concentration of free thiol remained. After 4.5 hours, only 2% of the beginning free thiol concentration remained. Biotinylation with PEO-iodoacetyl biotin at pH 8.50 was fast (half life ~20 min by DTNB) and very selective, producing a single species by RP-HPLC (Figure 10).

Biotinylation followed by avidin affinity chromatography was sufficient to remove the majority of unmodified CheY from the phosphonomethylation reaction mixture. Two passes over the avidin column were necessary, however, since each would only remove between 1 and 1.5 mg of biotinylated CheY. The resultant mixture was not always 100% phosphono-CheY. Side peaks as well as unmodified D57C CheY often remained (Figure 11a). Of the 23 phosphomethylation reactions performed, 20 were taken through all stages of purification producing 36.7 mg of purified *E. coli* phosphono-CheY. Of the 3 reactions that were not purified, one was lost during a concentration mix-up, and 2 were heterogeneous by RP-HPLC and did not warrant purification. Additionally, three samples were not quantified at the final stage of purification. Two samples containing 4 mg of the purest phosphono-CheY (Figure 11b) were sent to Noreen Francis at Brandeis University for visualization of the protein at the base of the flagellar motor by electron microscopy. No further results from this sample submission have been obtained.

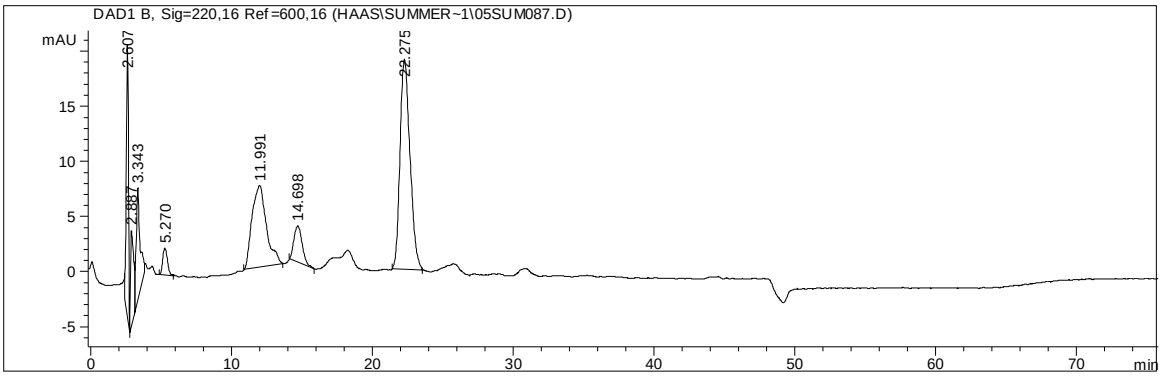
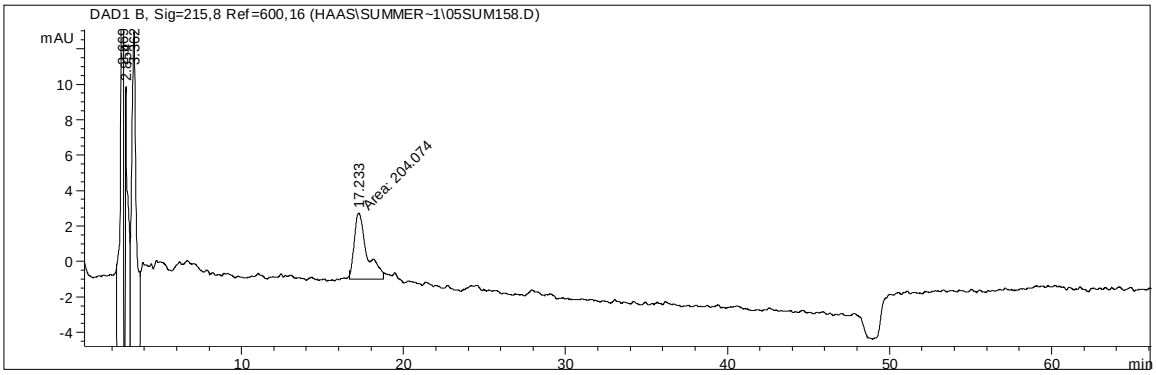
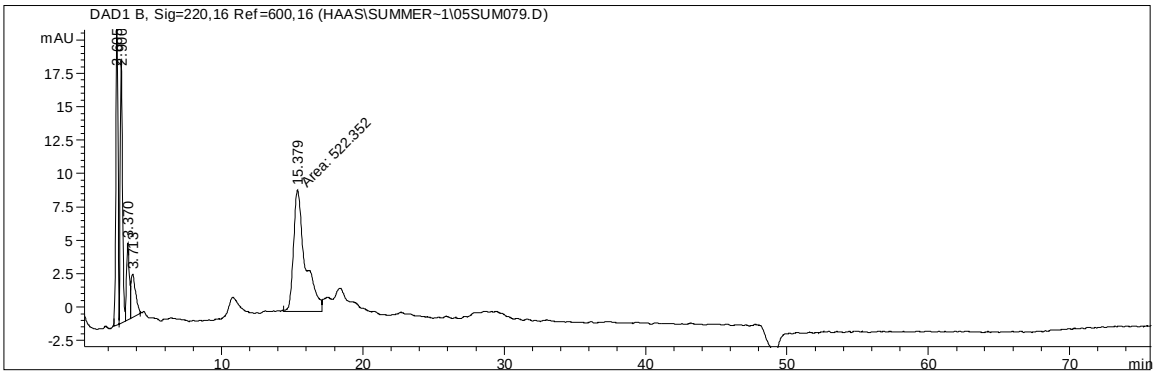


Figure 7. RP-HPLC chromatogram of S '05 J'. The 11 min peak is phosphono-CheY occupying 36% of the area. The 22 min peak is unmodified CheY. The 14 min peak is unknown.



a.



b.

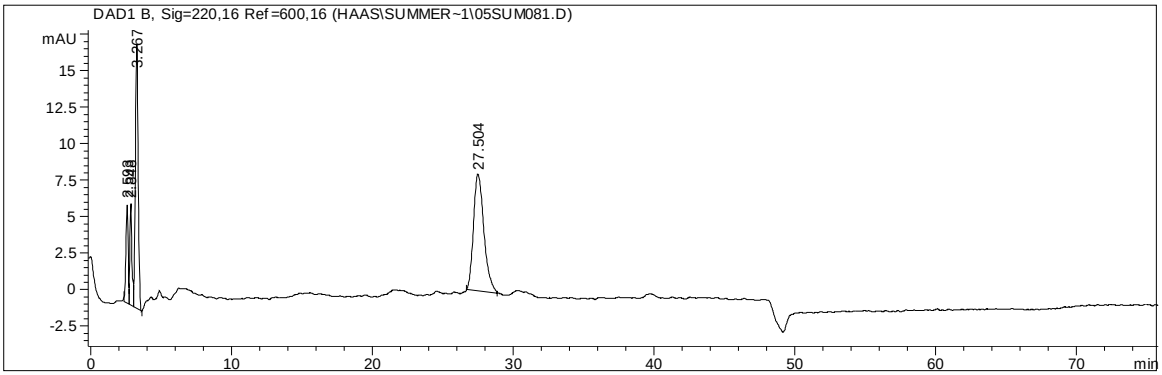
Figure 8. a.) RP-HPLC chromatogram of Sa '05 E'. The only species present is phosphono-CheY at 17 min. b.) RP-HPLC chromatogram of S '05 H'. The phosphono-CheY peak at 16 min occupies greater than 90% of the total area.

Phosphonomethylation	DTNB Δ % FT	RP-HPLC % modification
S '05 A	76	--
S '05 B	36	95
S '05 C	59	--
S '05 D	47	36
S '05 F	48	94
S '05 G	32	55
S '05 H	57	97
S '05 J	32	45
S '05 K	32	--
Sa'05 A	45	80
Sa'05 B	51	--
Sa'05 C	54	90
Sa'05 D	70	91
Sa'05 E	82	96
Sa'05 F	44	--
Sp'05 B	49	68
Sp'05 C	49	69
Sp'05 D	--	--
Sp'05 E	36	--
Sp'05 F	56	82
Sp'05 G	50	--
Sp'05 H	64	65
F'04 A	48	35

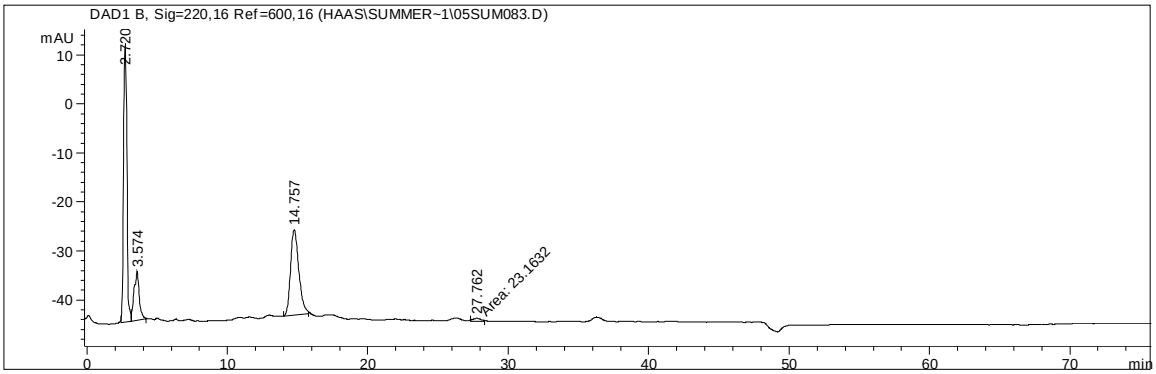
Table 1. Comparison of DTNB Percentages with RP-HPLC Alkylation Results. DTNB Δ % FT is the percent free thiol before reaction less the percent free thiol after modification. RP-HPLC % modification is the percent of the phosphono-CheY peak(s) area compared to the sum of the areas of the phosphono-CheY peak(s) and unmodified CheY peak(s) together. The percent free thiol before reaction of Sp'05 D was not obtained due to interference from excess reducing agent present. For reactions marked "--," no signal at 220 nm was obtained or no phosphonomethylation sample was analyzed by RP-HPLC.

Time (min)	[Free Thiol] (mg/mL)	CheY Peak Percentage
0	2.32	100
25	0.25	3
45	0.25	--
65	0.28	--
95	0.21	--
163	0.11	0
266	0.05	--

Table 2. Biotinylation of *E. coli* D57C CheY with PEO-Iodoacetyl Biotin, pH 8.5. DTNB indicates ~20 min half-life assuming pseudo-first order kinetics, and RP-HPLC suggests the reaction may be complete by 60 min. CheY peak percentage refers to unmodified D57C CheY and is calculated as the area of the CheY peak divided by the sum of the areas of the CheY and phosphono-CheY peaks.



a.



b.

Figure 9. a.) RP-HPLC of time 0 of the PEO-iodoacetyl biotinylation of *E. coli* D57C CheY. b.) RP-HPLC of time 25 min of the PEO-iodoacetyl biotinylation of *E. coli* D57C CheY. Unmodified CheY elutes at 28 min and CheY biotinylated with PEO-iodoacetyl biotin elutes at 15 min. Biotinylated CheY is essentially the only species present 25 min after the start of the reaction.

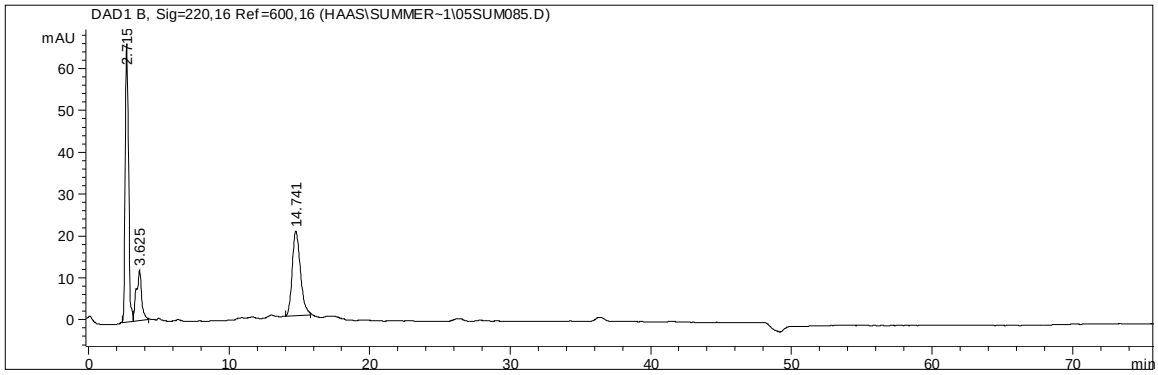
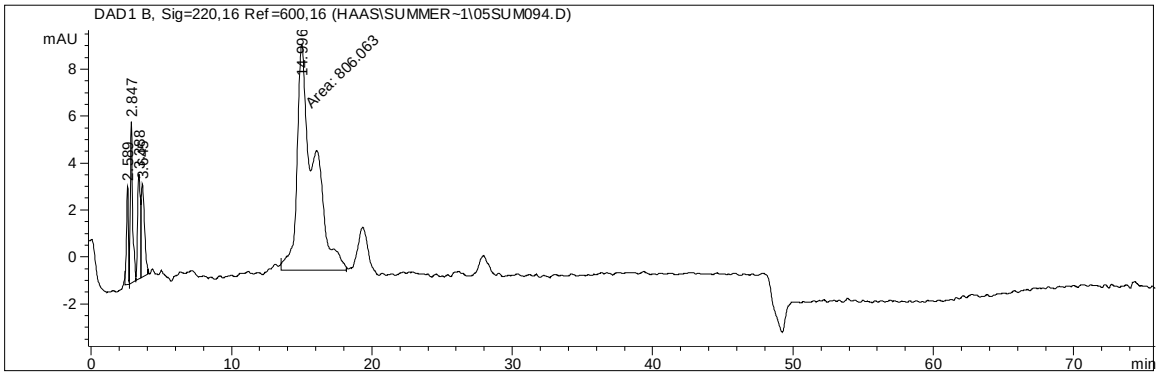
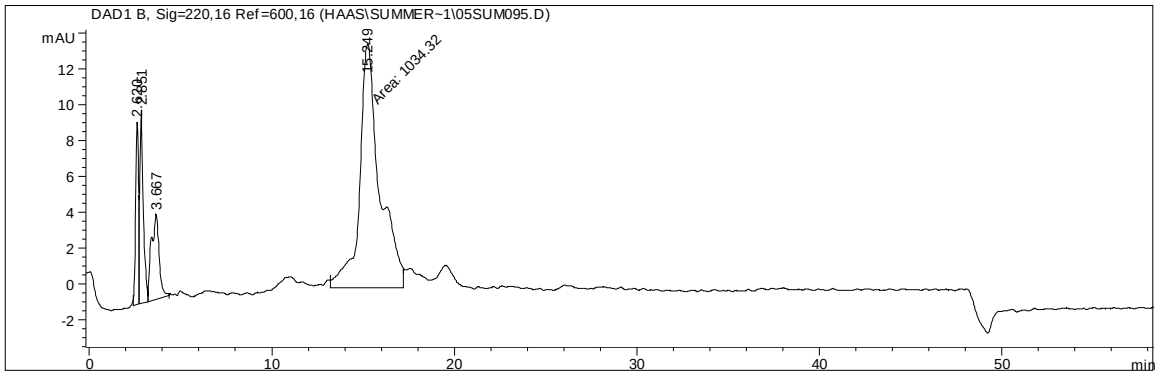


Figure 10. RP-HPLC of PEO-iodoacetyl biotinylated *E. coli* D57C CheY 8.25 hrs into the reaction. The peak at 15 min is biotinylated CheY.

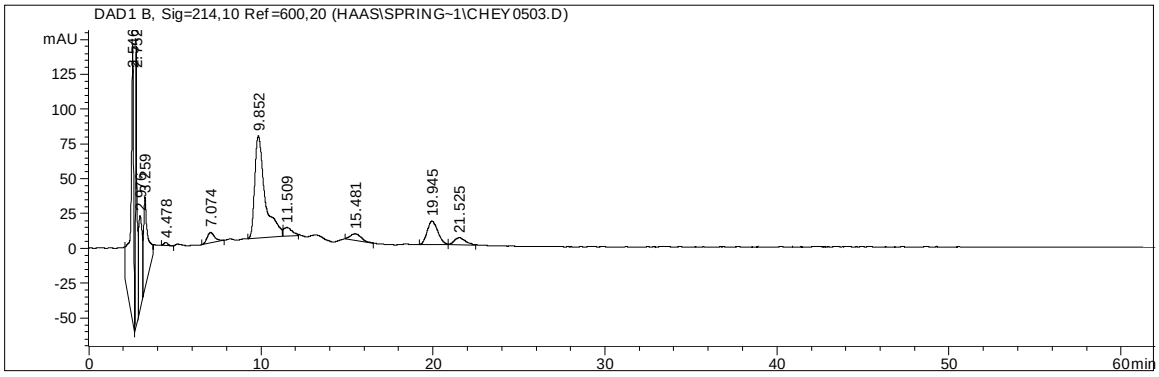


a.

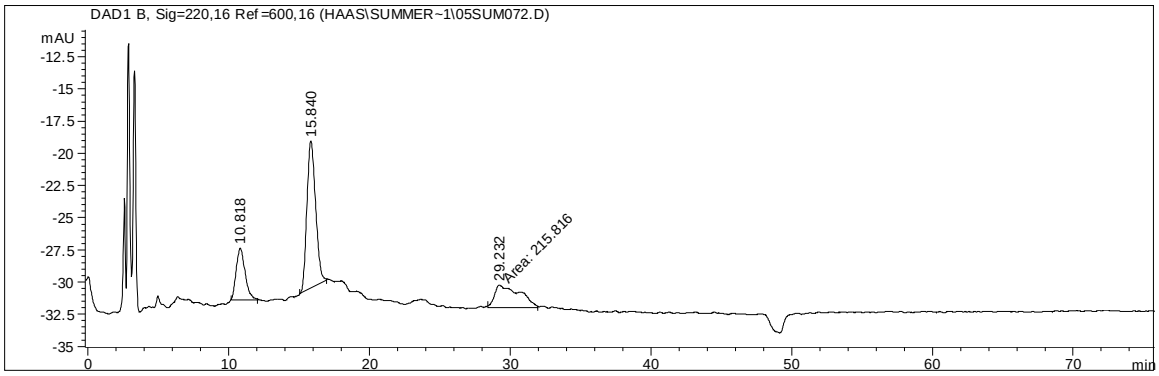


b.

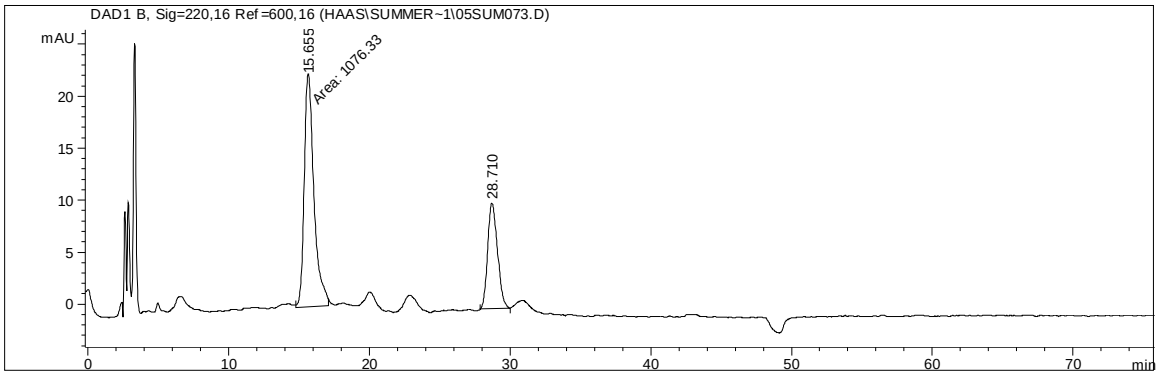
Figure 11. a.) RP-HPLC of partially purified *E. coli* phosphono-CheY. The phosphono-CheY peak elutes at 15 min. The peak on the right shoulder of the phosphono-CheY peak is unidentified but may be biotinylated CheY. b.) RP-HPLC of completely purified *E. coli* phosphono-CheY



a.



b.



c.

Figure 12. a.) RP-HPLC of the phosphonomethylation reaction mixture S ’05 F’ before being taken over the cibacron blue column. b.) RP-HPLC of the low salt wash of the cibacron blue column shows both phosphono-CheY (15 min) and unmodified D57C CheY (29 min) elute along with another species at 10 min that may be di-phosphonomethylated CheY. c.) RP-HPLC of the high salt (2M NaCl) wash of the cibacron blue column. Unmodified CheY has some affinity for the cibacron blue column since a greater proportion is present in the high salt wash.

Purification of Phosphono-CheY by Cibacron Blue Chromatography

A 3 mL cibacron (affigel) blue column was only partially successful as a means for purifying phosphono-CheY from the reaction mixture (Figure 12). Both phosphono-CheY and unmodified CheY were present in the low salt wash and the 2 M NaCl wash, however, the proportion of unmodified CheY in the high salt wash was greater than in the low salt wash, indicating a slight preference of unmodified CheY for the reactive dye. The fact that twice as much protein (2.5 mg) was found in the 2 M NaCl wash than in the low salt wash (1.2 mg) suggests that both phosphono-CheY and unmodified CheY show affinity for the dye. The only species in the mixture that was completely separated in the experiment eluted at 10.8 min and was not identified.

Production and Purification of *T. maritima* D54C/C81S CheY

Transformation by electroporation of the plasmid encoding *T. maritima* D54C/C81S CheY into *E. coli* cell lines BL21, B834 (DE3), and DH5-alpha was successful. Time and voltage constants for the electroporation of B834 (DE3) were 5.5 milliseconds and 2.50 kV, respectively. Comparable constants were seen in the transformation of BL21 and DH5-alpha. Colonies from each transformed cell mixture plated onto LB agar grew on the kanamycin plates but failed to grow on the ampicillin plates, indicating the bacteria had taken in and expressed the plasmid DNA. Cell line B834 (DE3) was used for production of the recombinant protein and the DH5-alpha cells were stored for future DNA plasmid purifications. A single colony of B834 (DE3) picked for a test growth showed IPTG-induced production of the recombinant protein by SDS-PAGE and coomassie staining and was frozen for future inoculations. Interestingly, even non-induced cells produced a protein band at or near the position of our recombinant CheY

protein (15.3 kDa), although it was less intense. Test inductions were also done to find the best time and [IPTG] at induction. 0.2 mM IPTG was used to induce 1 L of cells at an OD_{600} = 1.06 and 0.5 mM IPTG was used to induce 1 L of cells at an OD_{600} = 0.600. The growth induced with a lower [IPTG] later in the growth produced ~ 3% greater cell weight, however, both growths produced comparably intense CheY bands by 6.25 hours. All recombinant CheY was found in the soluble fraction based on SDS-PAGE.

The first preparations of *T. maritima* D54C/C81S CheY (M, N, O, R and T) were done according to the Cornell protocol with only a few minor additions and produced protein that appeared pure by SDS-PAGE but contained two species (presumably two different populations of CheY) that elute under the standard shallow RP-HPLC gradient defined in the HPLC Methods section (Figures 11-13). Notice that all of these preps had an impurity peak (CheY*) eluting ~27 minutes before CheY. Precipitation of protein was a common problem throughout preparations M-T, despite the presence of 2-ME or DTT. A flaky pale-yellow precipitate was found mainly during the dialysis steps and concentration steps, but would also form towards the end of a thrombin digestion. The protocol was manipulated to minimize precipitation due to dialysis and concentration, and this correlated with a more homogenous protein mixture when analyzed by RP-HPLC.

In preparations that showed only a single species by RP-HPLC, the cells were sonnicated with 1 mM PMSF and 10 mM 2-ME. The soluble fractions were loaded onto a 4 mL Ni-NTA column according to standard protocol, but 2-ME was left out completely during IMAC purification. The eluate was kept to between 5 – 10 mL, which was sufficient to remove all white color (bound protein) from the column. 5 mM $MgCl_2$, 2.5 mM $CaCl_2$, and 10 mM 2-ME was added directly to the eluate and digested with 1:2000 w:w thrombin:protein using high-

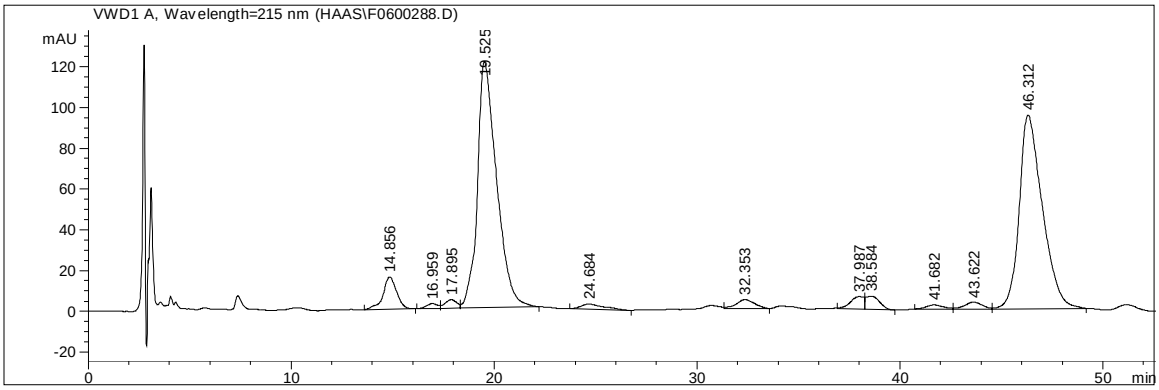
purity grade thrombin from MPBiomedicals. Once cleavage was confirmed by SDS-PAGE, 1 mM PMSF was used to stop thrombin digestion and the protein was passed through a 450 mL G-50 size-exclusion column that had been equilibrated with at least 1 liter of 50 mM phosphate buffer, pH 7.2 with 1 mM EDTA, 0.02 % azide, 0.5 mM PMSF, and 10 mM 2-ME. The same buffer was used to elute the protein, although PMSF was not always added. Fractions were analyzed spectrophotometrically by absorbance at 280 nm and three peaks were consistently found (Figure 5). The flow rate was kept to ~0.5 mL/min to achieve adequate separation. 50 mM 2-ME was added to fractions containing CheY and they were combined and concentrated to greater than 7 mg/mL. Precipitation was virtually eliminated even down to a protein concentration of 46 mg/mL. Protein prepared in this way provides a single major peak by RP-HPLC (Figure 16-15).

Protein preparations that included dialysis into thrombin cleavage buffer and more than one concentration step failed to produce a single species by RP-HPLC. These preparations were performed as follows. Cells were sonnicated in the presence of 1 mM PMSF and 1 mM DTT (except preps M and N.) The soluble fractions were passed over the Ni-NTA column and eluted in 7-12 mL of elute buffer. All buffers contained 5 mM 2-ME to keep the protein reduced throughout IMAC purification. A 1 mL Ni-NTA column was used for preparations M and N but a significant amount of recombinant CheY was found in the wash so subsequent preparations used a 4 ml column. The Ni-NTA column had a much greater binding capacity than that defined by the Novagen literature, binding ~100 mg of his-tagged protein per mL of settled gel. 95% purification of D54C/C81S CheY was achieved over the Ni-NTA column. The Ni-NTA wash of preparations M and N contained a significant amount of CheY and were treated as separate pools. The Ni-NTA eluate was dialyzed 1200x into thrombin cleavage buffer with 5 mM DTT.

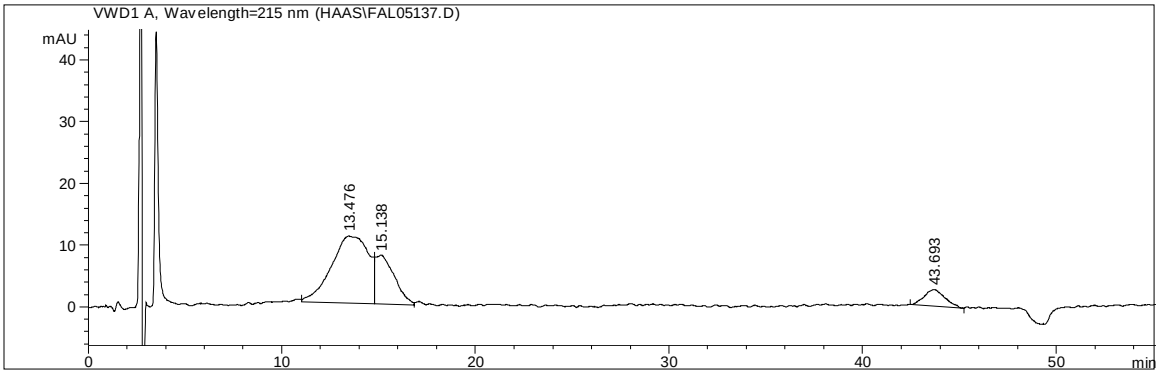
1 unit of restriction grade thrombin was used for ever 4 mg of protein and digestion reached completion by 18 hours. A digestion done with 0.20 units/mg protein was not complete at 18 hours. SDS-PAGE was sufficient to judge the completeness of thrombin digestion before the mixture was loaded onto the size exclusion column for the final purification step.

Pool M and pool N were passed through the gel filtration column into 50 mM Tris, pH 7.5 with 150 mM NaCl as described. Pool O was passed through the column into 50 mM MOPS, pH 7.0 with 1 mM EDTA. Pool R and T (and all subsequent preps) were passed through the column into 50 mM phosphate buffer, pH 7.2 with 1 mM EDTA. All gel filtration buffers contained 0.02% azide and 5 - 10 mM 2-ME. CheY was not completely separated from thrombin and other higher molecular weight impurities in the SEC if the flow-rate approached 1 mL/min. M pool 1 protein was not completely separated and as a result, M pool 2 contained the digested Ni-NTA wash as well as the sidecuts from the M pool 1 SEC fractions. N pool 2 also contained the sidecuts as well as the digested Ni-NTA wash. Pool 2 protein possessed no significant impurities and behaved the same on analytical reversed-phase HPLC compared to Pool 1 protein.

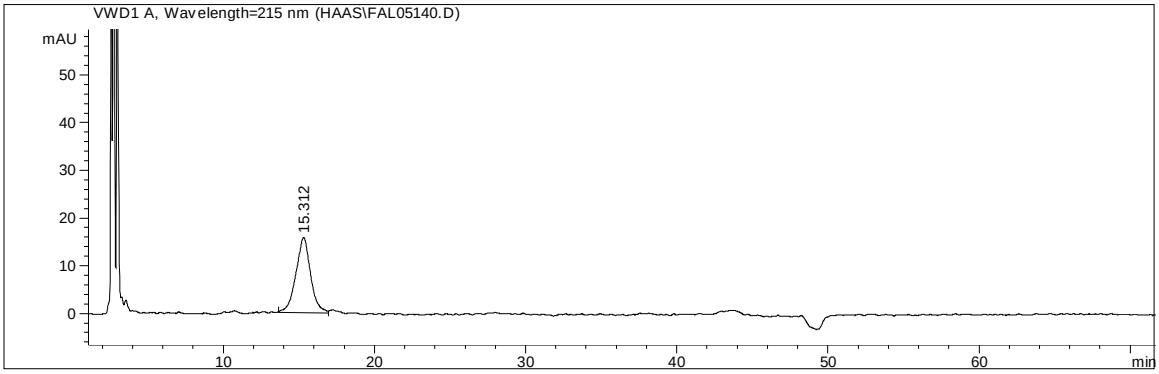
Size exclusion chromatography was successful as the final step in purification of *T. maritima* CheY. SDS-Page identified the first peak and second peaks as protein of molecular weight higher than 30 KDa and CheY, respectively. The third peak produced a streak of stain below 10 KDa that was thought to be the His-tag. A sample of this was combined with Ni-NTA resin, centrifuged, and the supernatant examined by gel. Protein was not present in the supernatant, and the peak was confirmed as the his-tag.



a.

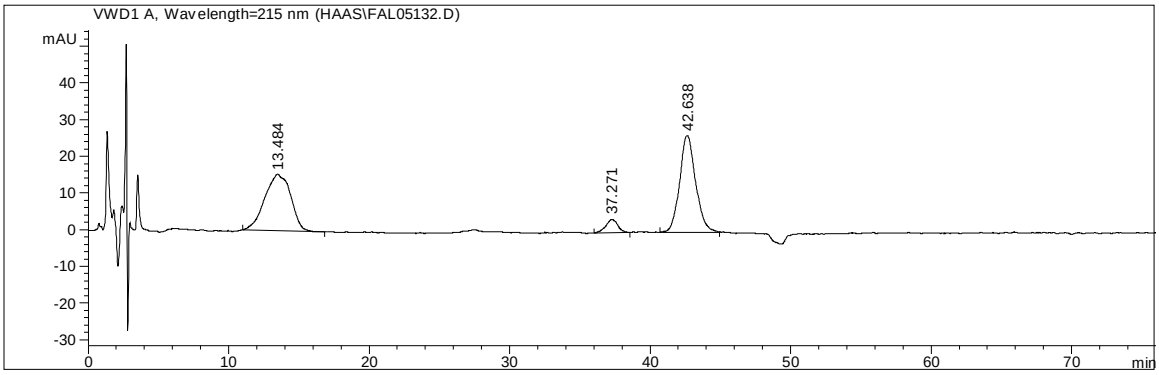


b.

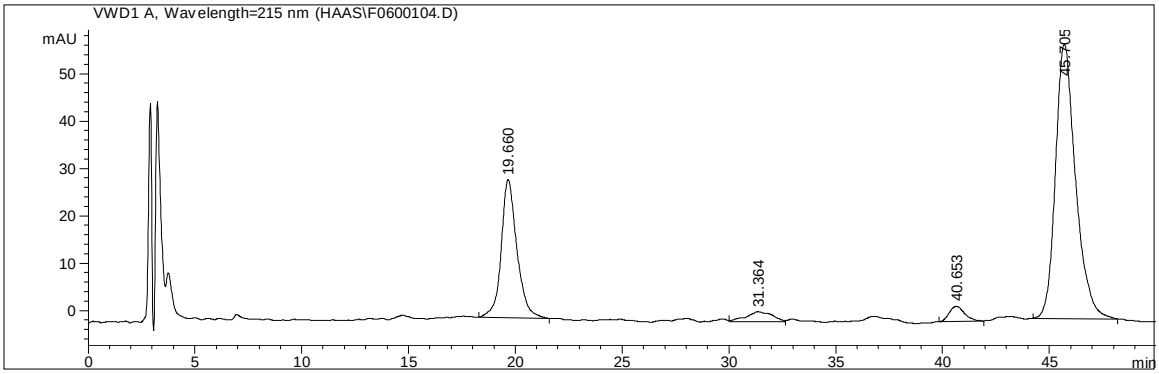


c.

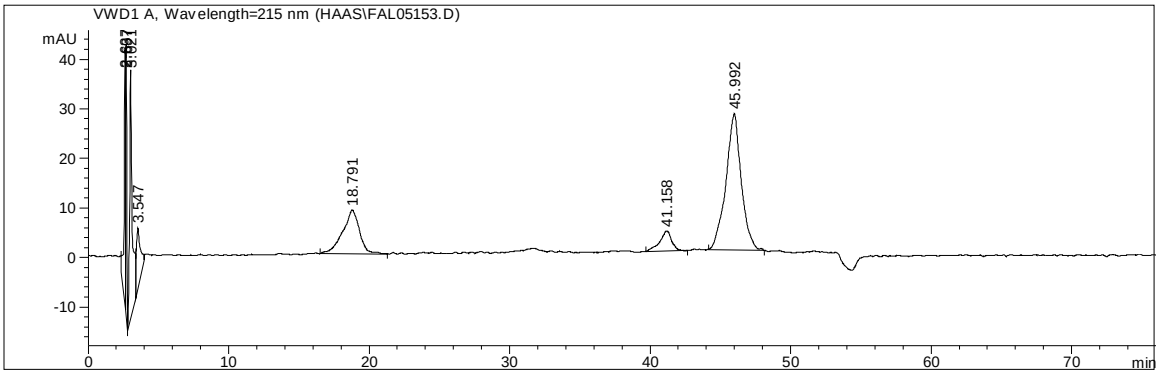
Figure 13. a.) RP-HPLC of preparation M pool 2 of *T. maritima* D54C/C81S CheY. CheY is the later eluting species. The large 20 min peak is an unknown species that is most likely some form of CheY and was given the title CheY*. b.) N pool 1 of *T. maritima* D54C/C81S CheY. c.) N pool 2 of *T. maritima* D54C/C81S CheY. CheY elutes at 44 min and CheY* at 15 min. Only CheY* is present in N pool 2.



a.

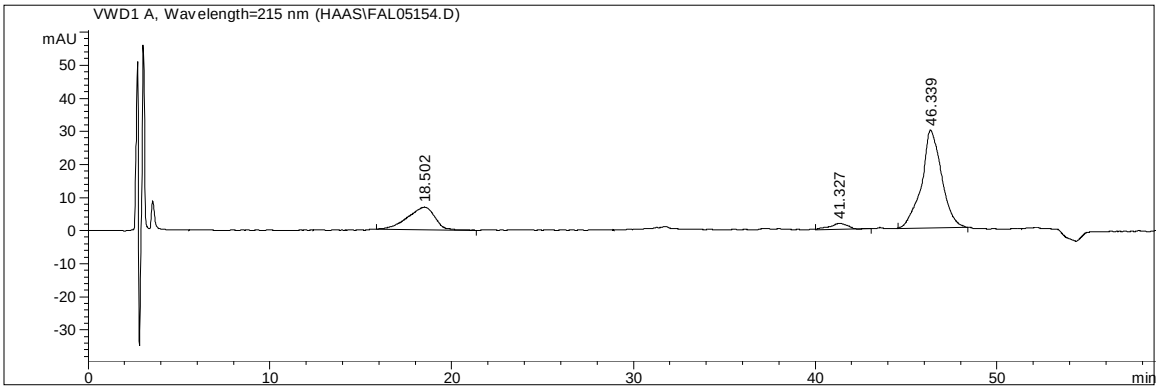


b.

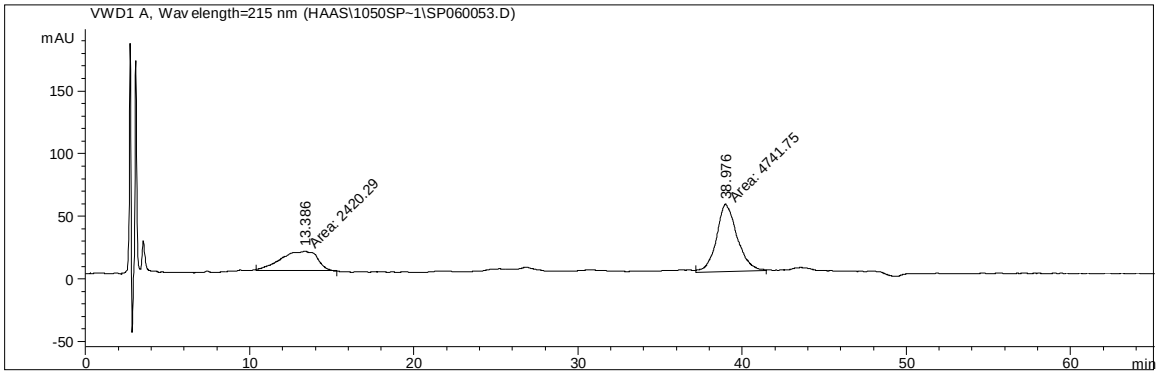


c.

Figure 14. a.) RP-HPLC of O pool 1 of *T. maritima* D54C/C81S CheY. b.) RP-HPLC of O pool 2 of *T. maritima* D54C/C81S CheY. c.) RP-HPLC of R pool 1 of *T. maritima* D54C/C81S CheY. CheY elutes between 43 and 46 min and CheY* elutes between 14 and 20 min. The differences in retention time are not believed to represent differences in the identity of the species but are a result of slightly different mobile phase composition.

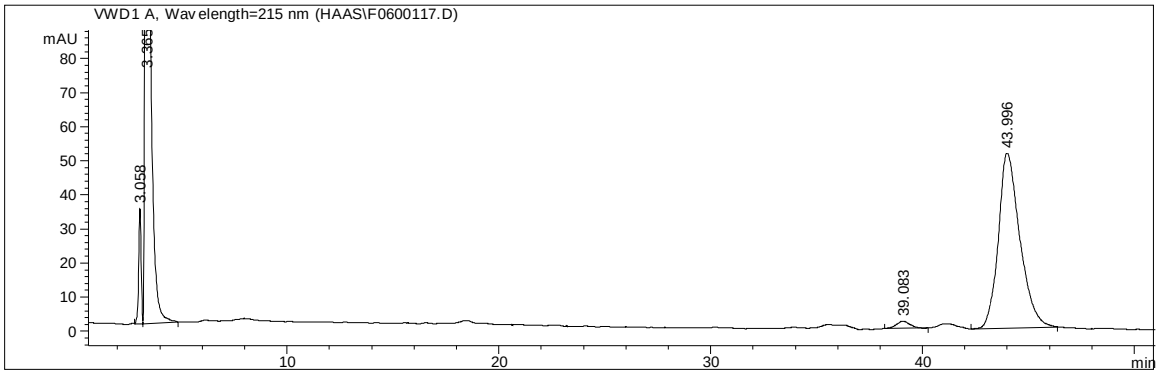


a.

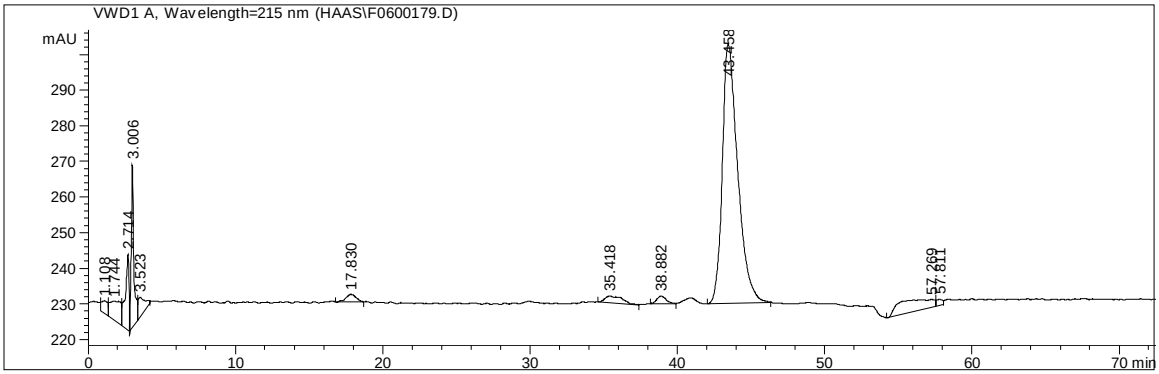


b.

Figure 15. a.) RP-HPLC of R pool 2 of *T. maritima* D54C/C81S CheY. b.) RP-HPLC of Prep T of *T. maritima* D54C/C81S CheY. CheY* elutes at 18 and 13 min and CheY elutes at 46 and 39 min. The peak at 41 min in panel a is oxidized CheY.

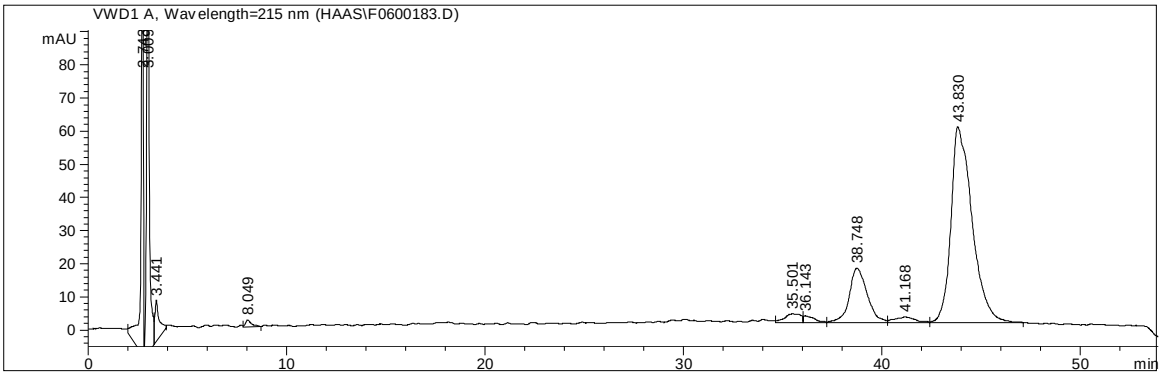


a.

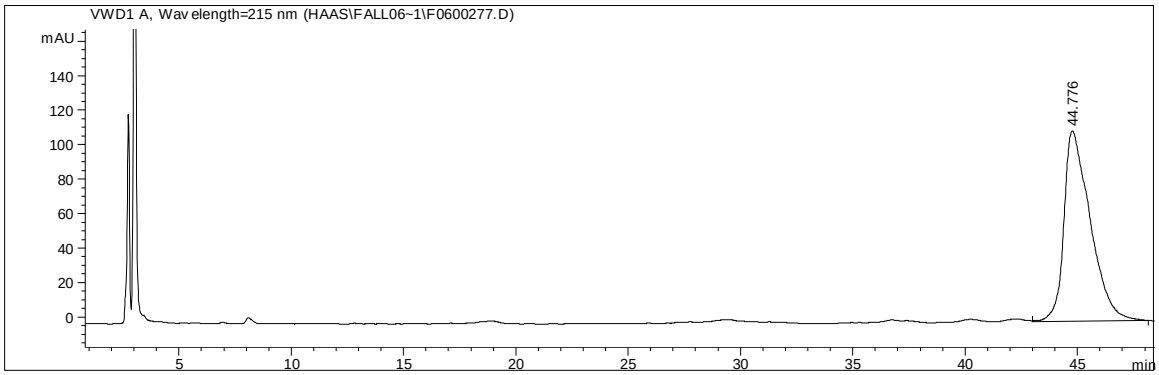


b.

Figure 16. a.) RP-HPLC of Prep A of *T. maritima* D54C/C81S CheY. b.) RP-HPLC of Prep B of *T. maritima* D54C/C81S CheY. One species elutes under the standard shallow gradient, presumably due to a single population of *T. maritima* D54C/C81S CheY.



a.



b.

Figure 17. a.) RP-HPLC of Prep C of *T. maritima* D54C/C81S CheY. b.) RP-HPLC of Prep D of *T. maritima* D54C/C81S CheY. One species elutes under the standard shallow gradient, presumably due to a single population of *T. maritima* D54C/C81S CheY.

Phosphonomethylation of *T. maritima* D54C/C81S CheY

Phosphonomethylation of *T. maritima* D54C/C81S CheY was carried out in a variety of conditions and produced on average 40-50% alkylation. Buffer concentration, pH, metal and metal concentration varied most often, but virtually every condition was tested and optimized around previous *E. coli* phosphonomethylation conditions. Concerns about the stability of the protein at high ionic strength and alkaline pH were also addressed before variables were changed. Results from *T. maritima* D54C/C81S CheY phosphonomethylations varied from approximately 20% to 90% by RP-HPLC peak areas and from 0% to 65% by DTNB assay. Calculation of percent alkylation from DTNB assays was done by subtracting the percentage of free thiol before and after alkylation. No correction was made for the percent of oxidized thiol present before reaction, which would increase apparent alkylation percentages. Calculation of percent alkylation by RP-HPLC peak area percentages were performed by dividing the area of the phosphono-CheY peak by the sum of the phosphono-CheY and unmodified CheY peak areas. Any peaks not presumed to be phosphono-CheY or D54C/C81S CheY were excluded from this calculation. The identity of the phosphono-CheY peak in RP-HPLC was based on previous results obtained from *E. coli* phosphonomethylations in which phosphono-CheY eluted ~18 minutes earlier than unmodified protein.

Variations of the phosphonomethylation conditions began by altering buffer strength and pH. TAPS at pH 8.50 was used at 250, 350, 425, and 500 mM concentrations. AMPSO at pH 9.0 was tried at the same four concentrations. Both buffers at 250 and 500 mM produced alkylations of greater than 50% completion, although TAPS was much more consistent when used at 250 mM. The reaction only reached 20-30% in 425 mM buffers unless the metal concentration was 125 mM or less at which point the alkylation percentage increased to near

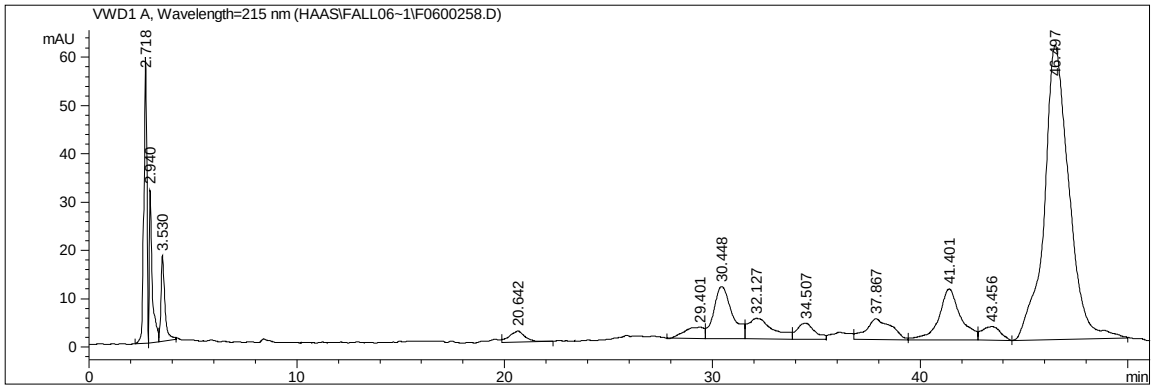
50%. Reactions done in 250 mM CAPSO at pH 9.6 and 10.1 produced less than 40% alkylation, but the mixture was fairly clean (Figure 18). 250 mM AMPSO, pH 9.25 did not produce better results. Modification of these higher pH reaction conditions may produce reactions of greater yield due to the increased nucleophilicity of the cysteine as alkalinity increases; however, the pK_a of the cysteine residue is unknown.

Metal concentration was also varied between 100 mM and 250 mM for the standard divalent metals $SrCl_2$, $BaCl_2$, and $CaCl_2$. In general all three metals produced equally good reactions. $SrCl_2$ produced alkylations of ~5% greater completion than $BaCl_2$ under the same conditions. $BaCl_2$ was more consistent though; 175 mM $BaCl_2$ in 250 mM TAPS, pH 8.50 regularly produced ~50% alkylation. $SrCl_2$ was never tried at 175 mM, but no consistency was seen with 200 mM $SrCl_2$. 125 mM $CaCl_2$ produced some of the best alkylations but also showed no consistency in reaction completeness. Trivalent metals produced impressive phosphonomethylation results. 10 mM $Nd(ClO_4)_3$ and $Lu(ClO_4)_3$ both produced alkylations better than 50% complete (Figure 19), however the resultant heterogeneity of the RP-HPLC analysis of the mixture indicates possible side reactions.

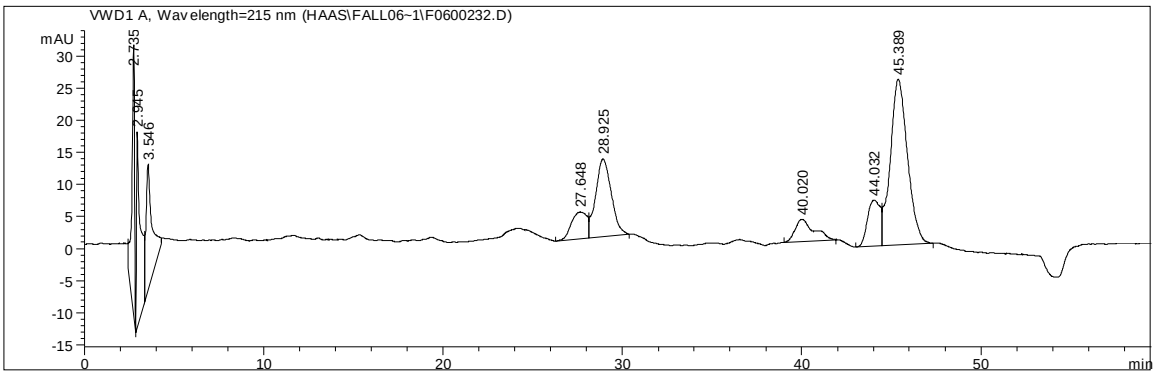
Triethylamine concentration also had an effect on alkylation percentage. In general, at the same protein concentration and with the same buffer and metal conditions, increasing the number of equivalents of triethylamine appeared to correspond with a slight increase in alkylation percentage. Figure 20 illustrates the ~10% increase in alkylation percentage as triethylamine equivalents were increased from 3 to 5. No more than 5 equivalents were ever used in an alkylation, though. Replacing the organic base with 2.8 equivalents of KOH produced alkylations of 50% or greater (Figure 21). AMPSO is used at concentration of 1.3 M and pH

10.1 for these alkylations. Lowering the metal concentration from 200 mM to 125 mM increased the alkylation percentage under these conditions.

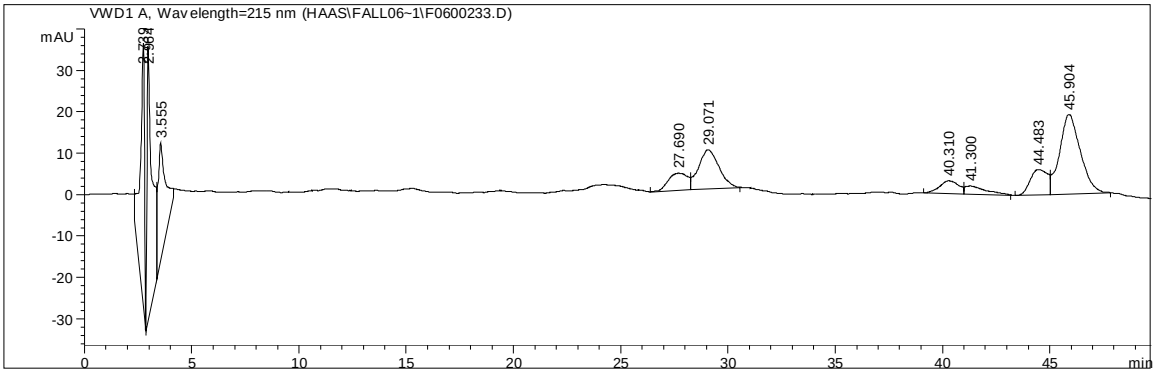
Variations in temperature and the batch of PMT used in phosphonomethylation had small effects on the resultant reaction mixture. Incubating the reaction at 37 °C seemed to produce slightly cleaner results by HPLC but did not increase reaction yield. A side-by-side comparison of C5' at room temperature and C6' at 37 °C illustrates the effect of increasing the temperature (Figure 22). Both reactions were done in 250 mM AMPSO, pH 9.25 with 125 mM $SrCl_2$ and 3 equivalents of triethylamine. A second alkylation (B5') under the elevated temperature gave the same result as C6' with only 5% greater yield. PMT 2006.AE, 2006.BE, 2005.C, and 2003.3 were each tested under exactly the same set of conditions: protein at 9.3 mg/mL in 250 mM TAPS, pH 8.50 with 175 mM $BaCl_2$ and 3 equivalents of triethylamine. Reactions with 2006.AE and 2005.C produced the greatest alkylation percentage by DTNB, however they were the poorest alkylations by RP-HPLC peak areas. 2006.BE gave the greatest percentage of alkylation by RP-HPLC (Figure 23). All of these chromatograms contained multiple peaks that did not correspond to the retention time of phosphono-CheY (or CheY), and data gathered by HPLC peak areas must be examined skeptically.



a.

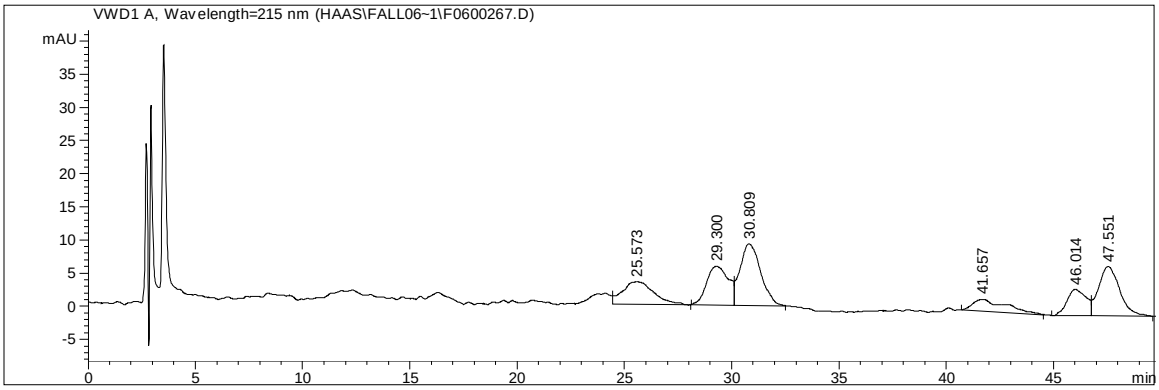


b.

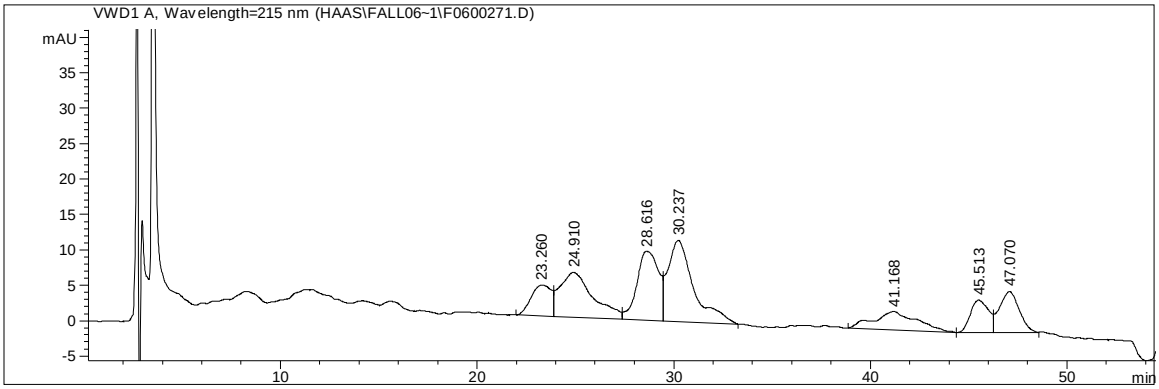


c.

Figure 18. a.) RP-HPLC of the phosphonomethylation of *T. maritima* D54C/C81S CheY in CAPSO, pH 9.6. The reaction was less than 20% complete but was fairly clean. The 31 min peak is phosphono-CheY and the 46 min peak is unmodified CheY. Oxidized CheY elutes at 41 min. b.) RP-HPLC of the phosphonomethylation of *T. maritima* CheY in CAPSO, pH 9.6 indicates a clean 40% alkylation. c.) RP-HPLC of the phosphonomethylation of *T. maritima* CheY in CAPSO, pH 10.1 shows another clean 40% alkylation.

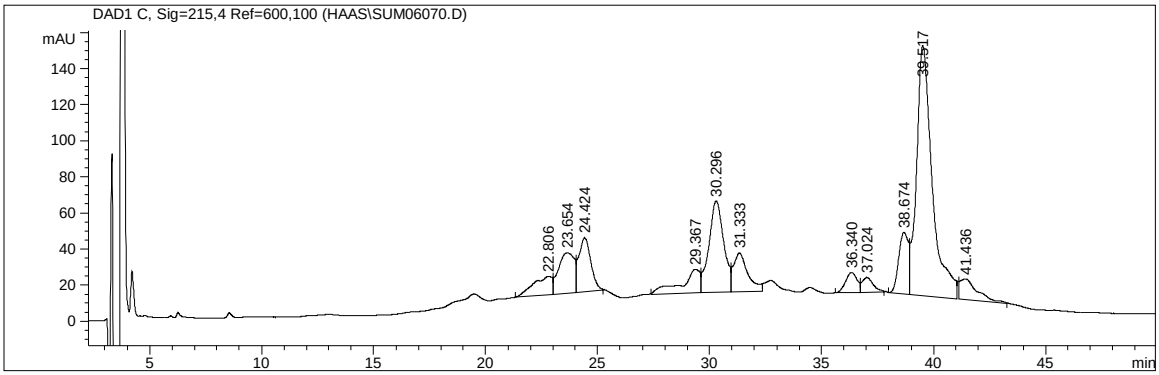


a.

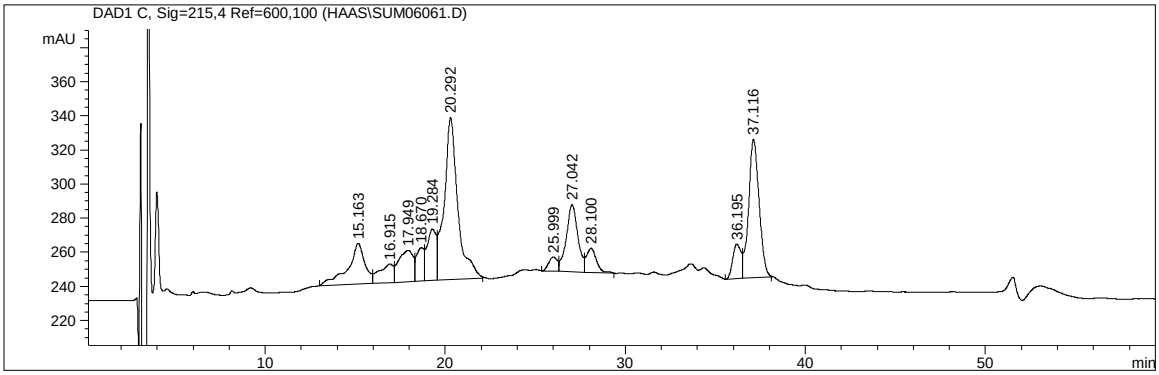


b.

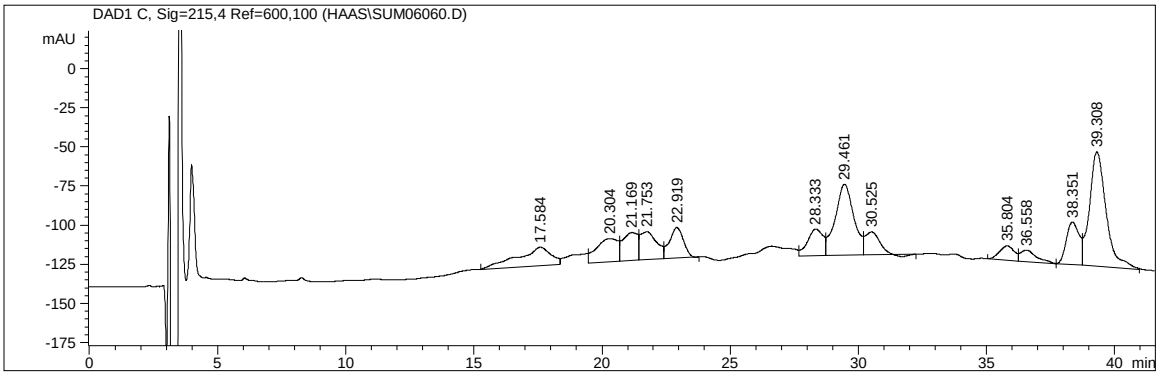
Figure 19. a.) RP-HPLC of the phosphonomethylation of *T. maritima* D54C/C81S CheY in 250 mM AMPSO, pH 9.25 with 10 mM $Nd(ClO_4)_3$. The reaction is greater than 50% complete but a peak at 25 min has appeared. B.) RP-HPLC of the phosphonomethylation of *T. maritima* D54C/C81S CheY in 250 mM AMPSO, pH 9.25 with 10 mM $Lu(ClO_4)_3$. The reaction is not clean but is almost 75% complete.



a.

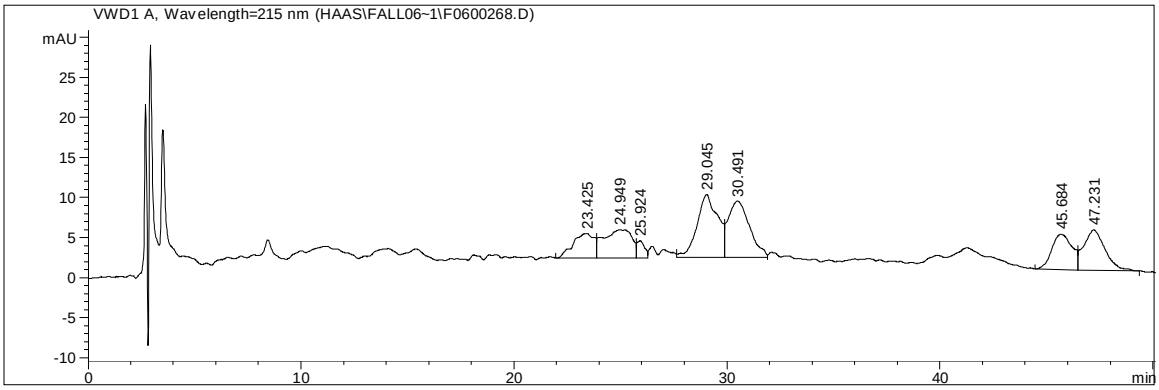


b.

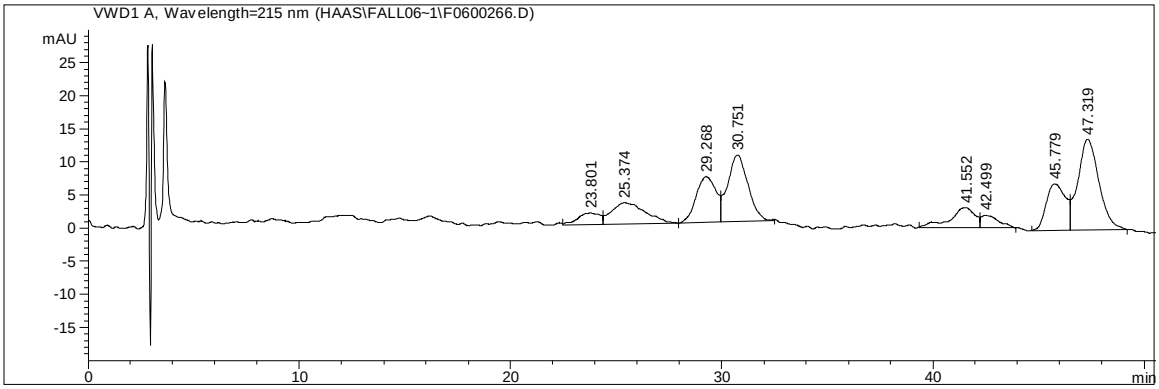


c.

Figure 20. RP-HPLC of the phosphonomethylation of *T. maritima* D54C/C81S CheY in 250 mM AMPSO, pH 9.0 with 3 (a), 4 (b), and 5 (c) equivalents of triethylamine. Phosphono-CheY elutes as a triplet near 28 min and unmodified CheY elutes near 38 min. Reaction percentage increases ~5% with the addition of each equivalent of triethylamine. A large amount of heterogeneity exists in these reaction mixtures.

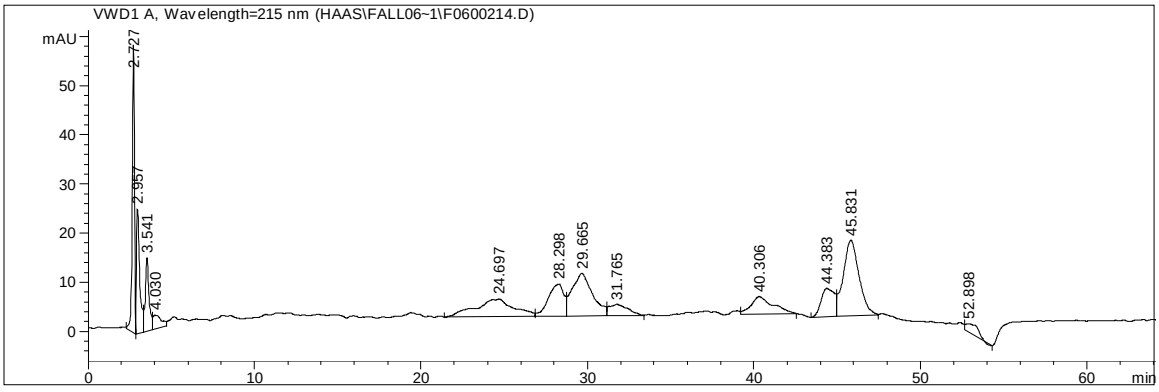


a.

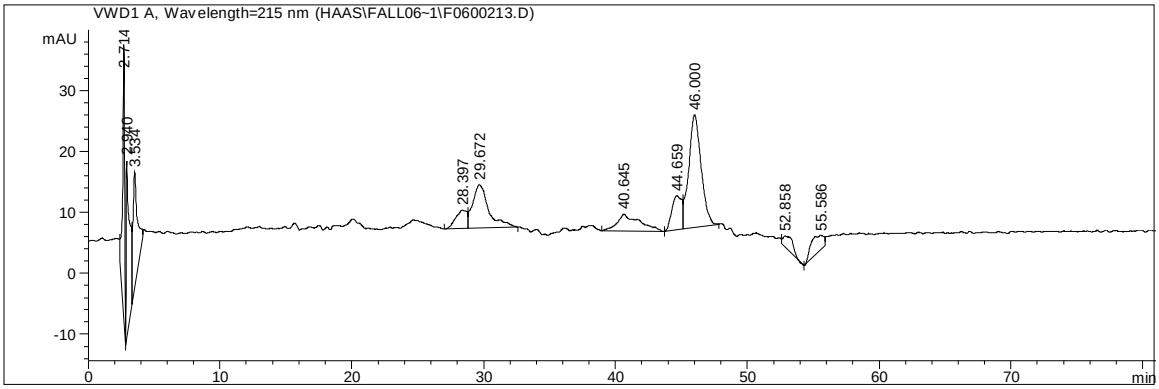


b.

Figure 21. a.) RP-HPLC of the phosphonomethylation of *T. maritima* D54C/C81S CheY in 1.3 M AMPSO, pH 10.1 with 125 mM $SrCl_2$ and 2.8 equivalents of KOH. b.) RP-HPLC of the phosphonomethylation of *T. maritima* D54C/C81S CheY in 1.3 M AMPSO, pH 10.1 with 200 mM $SrCl_2$ and 2.8 equivalents of KOH. Phosphono-CheY elutes as the doublet at 30 min and unmodified CheY elutes as the doubled at 46 min. The conditions produce good phosphonomethylation percentages.



a.



b.

Figure 22. a.) RP-HPLC of the phosphonomethylation reaction of *T. maritima* D54C/C81S CheY at room temperature (22°C). b.) RP-HPLC of the same phosphonomethylation reaction performed at 37 °C. The percentage of phosphono-CheY (29 min) formed drops slightly under elevated temperatures but may produce a slightly cleaner product.

Reaction	DTNB % Alk.	HPLC % Alk.
TMY1	17	50
TMY2	--	10
TMY3	--	10
TMY4	8	50
TMY5	13	35
TMY6	**42**	**40**
TMY7	**23**	**25**
TMY8	--	--
TMY9	41	65
TMY10	65	55
TMY11	36	50
TMY12	**63**	**60**
TMY13	48	40
TMY14	55	75

Table 3. Percentage alkylation of *T. maritima* D54C/C81S CheY by DTNB and RP-HPLC agreed for only 3 of the 14 reactions listed here. In general, RP-HPLC showed a greater percentage of alkylation than DTNB.

Purification of *T. maritima* Phosphono-CheY by Biotinylation/Avidin

Purification of phosphono-CheY from a reaction mixture by means of the biotinylation/avidin protocol was met with limited success, but a few putatively pure phosphono-CheY samples exist. Three biotinylation reagents were used, but prior to their use in purification of phosphono-CheY, the kinetics of the reactions between unmodified *T. maritima* D54C/C81S CheY and each reagent were determined. Table 4 shows the disappearance of CheY in terms of the remaining free thiol and RP-HPLC peak area percentage. The reaction with PEO-iodoacetyl biotin at pH 8.5 was faster than with PEO-maleimide biotin at pH 7.0, having an estimated half life of less than 20 min by RP HPLC compared to ~60 min for maleimide. DTNB gives data contrary to this; however, at the 0.5 hr time-point of the maleimide reaction, the DTNB absorbance at 410 nm was initially higher than usual and may have contributed to the smaller ΔA and hence a smaller concentration of free thiol. Biotin-HPDP was comparable in rate to PEO-iodoacetyl biotin.

Reaction of CheY with biotin-HPDP was followed spectrophotometrically by absorbance at 343 nm of the pyridine-2-thione leaving group, providing a unique opportunity to analyze the pseudo-first-order reaction in detail. Figure 24 shows a progress curve of this reaction. Using the first-order rate equation

$$\ln (A_\infty - A_t) = -kt + \ln (A_\infty - A_0)$$

gives the linear least-squares regression equation y = -6.56E-04x - 1.20 with $R^2 = 1.00$. From this, the rate constant k = 6.56E-04 and the half life (defined by first-order kinetics, $t_{1/2} = (\ln 2)/k$) is 1060 sec (17.6 min). A_∞ was defined as 0.384 and A_0 as 0.58 for this treatment. Using the logarithmic form of the first-order rate equation

$$A_t = -(A_\infty - A_0)*e^{-kt} + A_\infty$$

and allowing A_∞ and k to be variables initially defined as 0.385 and 0.000600, respectively, gives a fitted $A_\infty = 0.344$ and $k = 0.00117$. A half life of 591 seconds (9.5 min) is obtained from this value of k. Alternate initial values for A_∞ and k did not change the fitted values upon which ProStat converged.

The two methods above rely on an accurate determination of the starting point and endpoint of the reaction; however, the endpoint of the reaction with biotin-HPDP was difficult to determine due to a slow but measurable decay of the product. For this reason, the data were also analyzed by the Guggenheim method, a set of equations that can be used to determine first-order rate information when A_∞ is not known. Data values at 100 second intervals from 72 seconds to 972 seconds were subtracted from the values at t+1000 seconds (1072-72, 1172-172,1272-272, …1972-972) and put into the equation:

$$\ln(x_t - x_\Delta) = -kt + \ln(x_0 - x_\infty)(1 - e^{-k\Delta})$$

where x_t = Abs at t and x_Δ = Abs at t+1000.

A plot of $\ln (x_t - x_\Delta)$ vs. t for the 72 – 972 seconds gives a linear regression equation of $y = -0.00120x - 1.68$ with $R^2 = 0.967$. The rate constant $k = 0.00120\ s^{-1}$, giving a half life $\ln(2)/k$ = 578 seconds (9.6 min). Extending the time range to include 72 – 2072 seconds gives a linear regression of $y = -0.000900x - 1.80$ with $R^2 = 0.979$. The rate constant $k = 9.06E\text{-}04\ s^{-1}$, giving a half life of 765 seconds (12.8 min). A plot of $\ln (x_t - x_\Delta)$ vs. t for the first 1972 seconds gives a linear regression equation of $y = -0.00120x - 1.675$ with $R^2 = 0.967$. The rate constant k = $0.00120\ s^{-1}$, giving a half life $\ln(2)/k$ = 578 seconds (9.6 min).

All three reagents were used to biotinylate unmodified CheY before purification of phosphono-CheY. Characterization of CheY biotinylated with each of the three reagents was carried out prior to their use with a phosphonomethylation reaction mixture. Biotinylation with

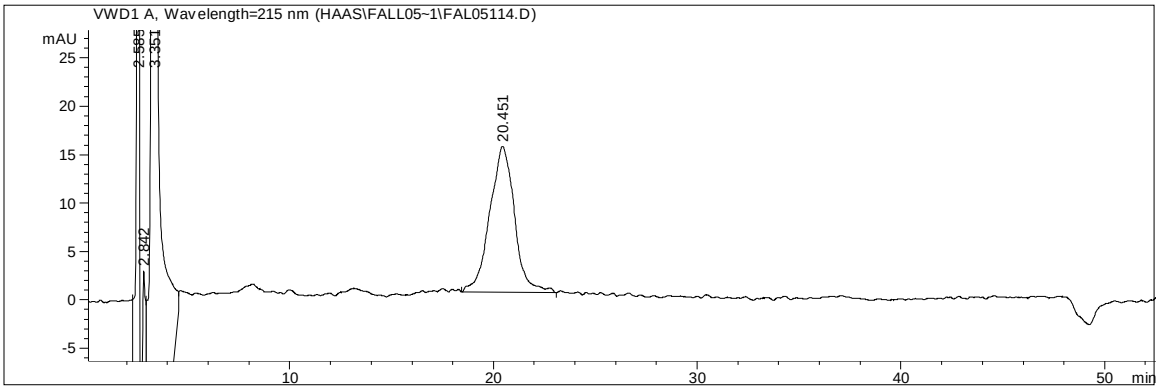
PEO-iodoacetyl biotin and PEO-maleimide biotin consistently removed the unmodified CheY peak in RP-HPLC and dropped the concentration of free thiol to near zero. The disadvantage of these two reagents is that the retention times of PEO-iodoacetyl and PEO-maleimide biotinylated CheY overlap with the phosphono-CheY peak in RP-HPLC (Figure 25), so determining the extent of reaction with these two reagents was difficult and relied on the disappearance of the CheY peak in RP-HPLC and minimal DTNB reactivity. Biotin-HPDP as advantageous in that respect; biotinylation with biotin-HPDP produced a species of CheY that eluted between phosphono-CheY and unmodified CheY by RP-HPLC, providing evidence for the presence of biotinylated CheY (Figure 25). Another advantage of using biotin-HPDP is the ability to monitor the progress of the reaction spectrophotometrically by an increase in A_{343}.

Biotinylation of unmodified *T. maritima* CheY was successful with all three reagents; however, immobilized monomeric avidin often failed to remove biotinylated-CheY from phosphono-CheY within a reaction mixture. This was most apparent when unmodified CheY was labeled with biotin-HPDP, since the biotinylated-CheY peak is visible in RP-HPLC analysis (Figure 26). This was less apparent when the water soluble reagents PEO-iodoacetyl and PEO-maleimide biotin were used since the retention times of CheY biotinylated with these reagents overlap that of phosphono-CheY in RP-HPLC analysis. A slight difference in the shape of the 30 min peak, presumed phosphono-CheY and biotinylated CheY co-eluting, was occasionally seen suggesting some sort of chemistry had occurred; however, the difference in peak shape was not reproducible from sample to sample and no conclusions could be made solely based on this evidence. On the other hand, the amount of protein present after the avidin column was reproducibly consistent with the proportion of biotinylated CheY present in the mixture, suggesting biotinylated CheY had been removed.

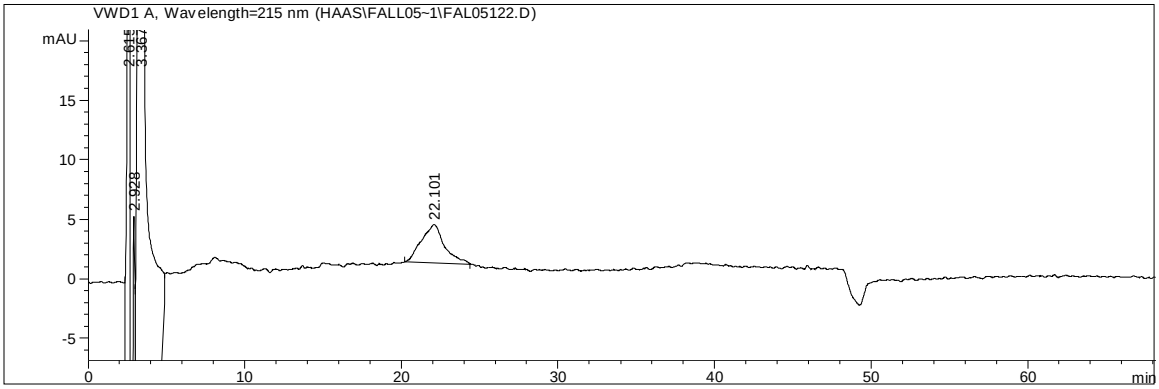
Reversed-phase analysis of the purification of phosphono-CheY from a reaction mixture was complicated by the apparent disagreement in evidence. Results of protein quantitation suggested removal of biotinylated CheY and adequate purification of a number of phosphono-CheY samples, but this was not the case when the purified triple-prime sample was analyzed by RP-HPLC (Figure 27). Continually, a species eluted in RP-HPLC at the same time as unmodified CheY. Only a few explanations can be found for this phenomenon. The first and least plausible is that the thioether bond formed in the reaction of PEO-iodoacetyl biotin and PEO-maleimide biotin with CheY is labile. There is no chemical precedent in the literature for a labile thioether bond. The second explanation is carry-over from previous analyses, possibly a result of precipitation of CheY on the column causing it to elute in the next reversed-phase experiment. Blank runs suggest that this is not the case however. The last explanation is that it is a new chemical species altogether. Spiking the sample with unmodified CheY would confirm this hypothesis if two peaks elute at 40 min, but definitive evidence would have to come from mass spectrometry analysis.

Results obtained from RP-HPLC analysis of the purification of phosphono-CheY were hard to interpret not only because of the apparent re-emergence of putative unmodified CheY after avidin, but also because of an unexplainable lack of signal obtained by double-prime samples. Analysis of a protein mixture after a biotinylation reaction (double-prime sample) consistently showed a decrease in peak area per µg of protein loaded. Presumed phosphono-CheY absorbs approximately 2.5-fold greater in the single-prime sample than in the double-prime sample even though the same amount of protein was loaded onto the column (Figure 28). Broad gradient reversed-phase HPLC analysis ruled out the possibility that protein species were eluting before the start of the shallow gradient. Desalting and buffer exchange also had no effect

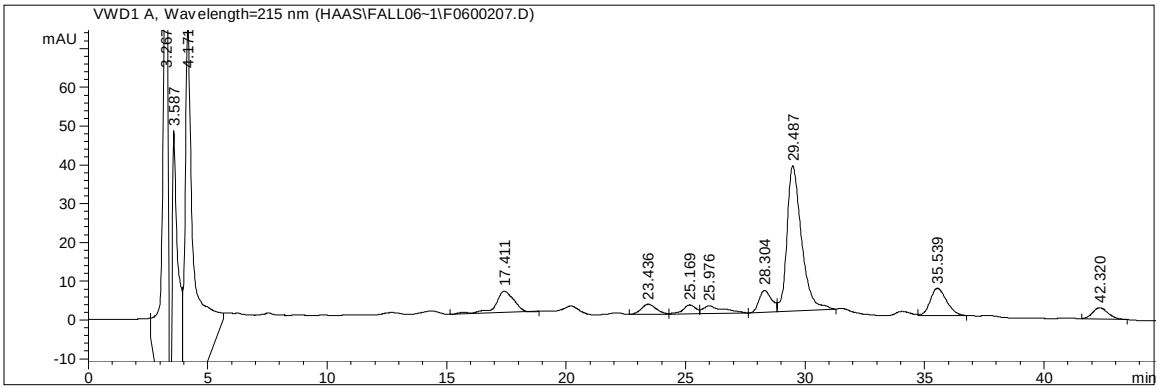
on the signal obtained. Interestingly though, this phenomenon was specific to the use of the two water soluble PEO biotin reagents. CheY modified with biotin-HPDP did not behave in this manner; signal remained proportional to the concentration of protein loaded. Biotin-HPDP seemed to shift the retention time of phosphono-CheY ~10 minutes earlier, though (Figure 29). The most obvious cause of this is chromatographic error, possibly in mobile phase composition or in sample composition. However, the retention time of HPDP-biotinylated CheY does not shift in a similar manner, providing evidence contradictory to chromatographic error. The shift in phosphono-CheY retention is possibly be the result of a non-specific interaction between phosphono-CheY and biotinylated CheY in which the biotinylated species disrupts the hydrophobic foot of phosphono-CheY and interferes with its binding to the stationary phase resulting in earlier elution. Despite these apparent shifts in reversed-phase elution, biotin-HPDP remains the reagent of choice since it can be followed spectrophotometrically, and it produces an identifiable peak that is distinct from phosphono-CheY and unmodified CheY in RP-HPLC.



a.



b.



c.

Figure 23. RP-HPLC of purified phosphono-CheY from *T. maritima.* The samples in panels a and b had to be biotinylated twice (once with PEO-iodoacetyl biotin and once with PEO-maleimide biotin) and taken over avidin twice to remove residual unmodified D54C/C81S CheY. Phosphono-CheY elutes at 22 min in panels and b. The sample in panel c was biotinylated with biotin-HPDP and had to be taken over avidin twice to remove residual biotinylated CheY. Phosphono-CheY elutes at 29 min, biotinylated CheY elutes at 36 min, and unmodified CheY elutes at 42 min.

	PEO-Iodoacetyl, pH 8.5		PEO-Maleimide, pH 7.0		
Time (hr)	**DTNB [FT] mg/mL**	**HPLC CheY peak % Area**	**Time (hr)**	**DTNB [FT] mg/mL**	**HPLC CheY peak % Area**
0	2.62 mg/mL	100 %	0	2.33 mg/mL	100 %
0.5	0.52 mg/mL	6 %	0.5	0.33 mg/mL	60 %
8.25	---	0 %	2.5	---	15 %
			4.5	---	9 %
			8.25	---	6 %

Table 4. Kinetics of biotinylation of *T. maritima* D54C/C81S CheY. The reaction with PEO-iodoacetyl biotin is much faster than PEO-maleimide biotin based on the disappearance of unmodified CheY in RP-HPLC analysis. Notice the disagreement in the two analytical methods at the 0.5 hr time-point of the maleimide reaction. This is possibly the result of a higher than normal absorbance at 410 nm to begin the assay, leading to a calculation of a lower percentage of free thiol.

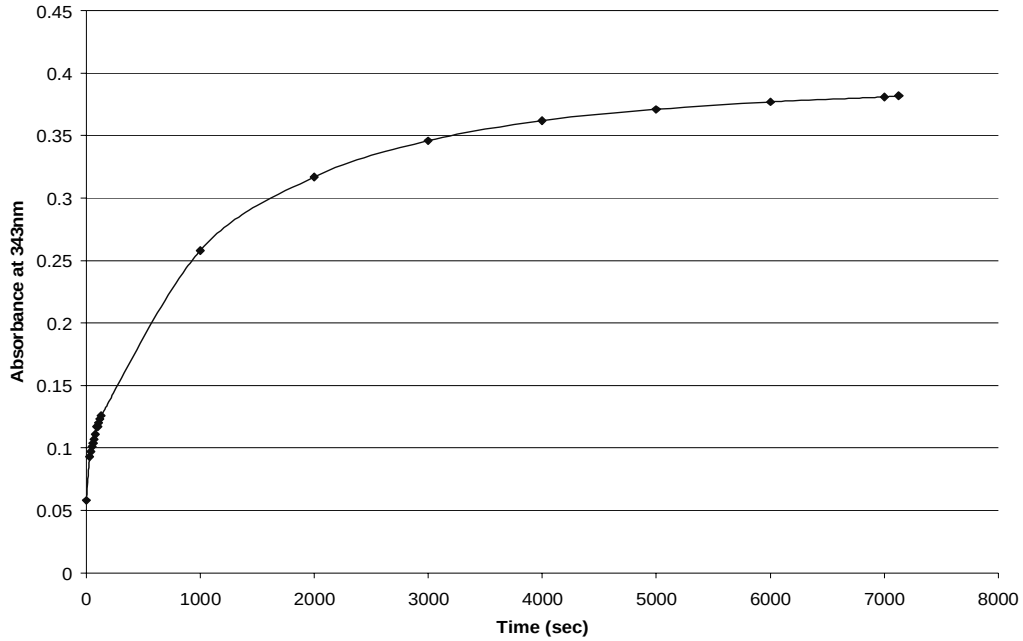
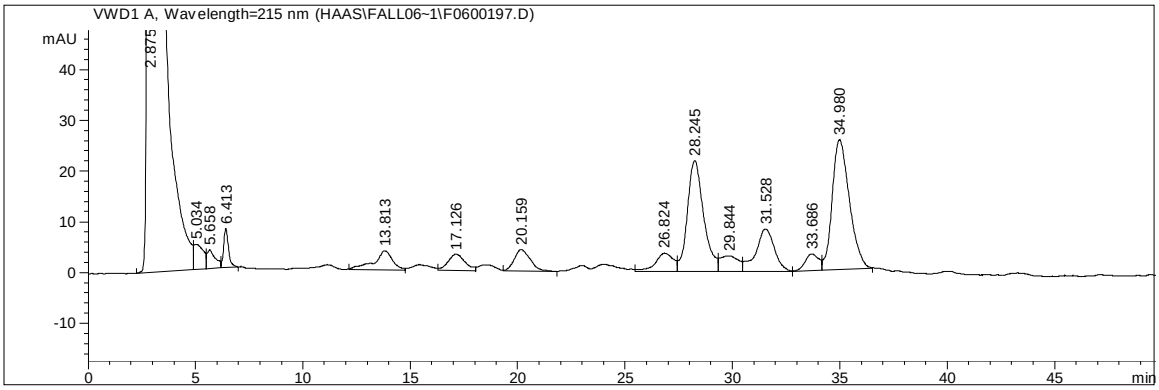
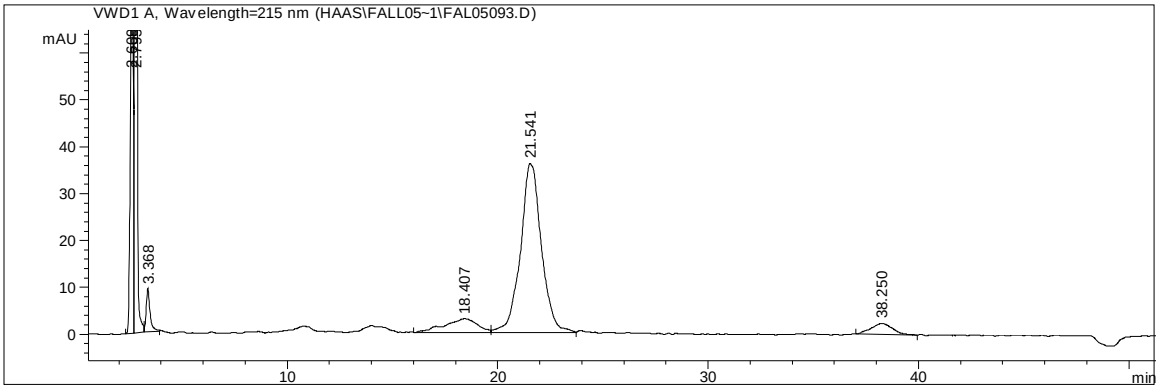


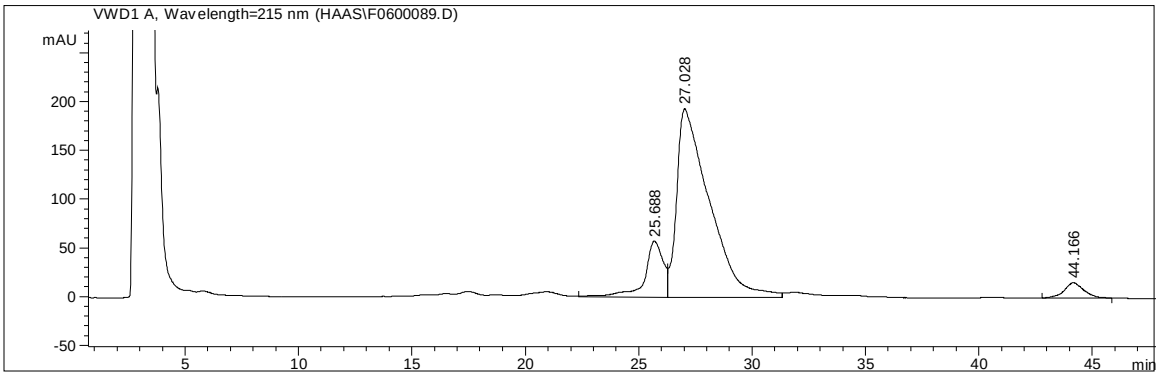
Figure 24. Progress curve of the reaction of *T. maritima* D54C/C81S CheY with Biotin-HPDP. The curve was produced by connecting data points measured spectrophotometrically.



a.

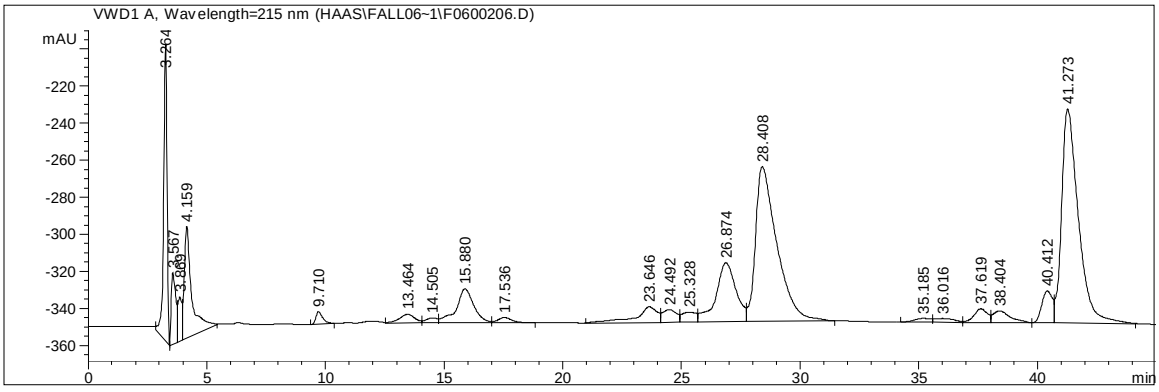


b.

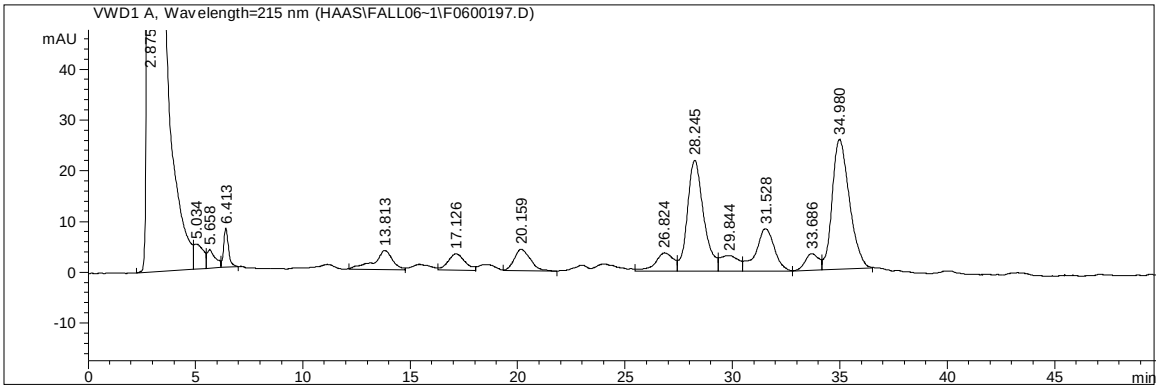


c.

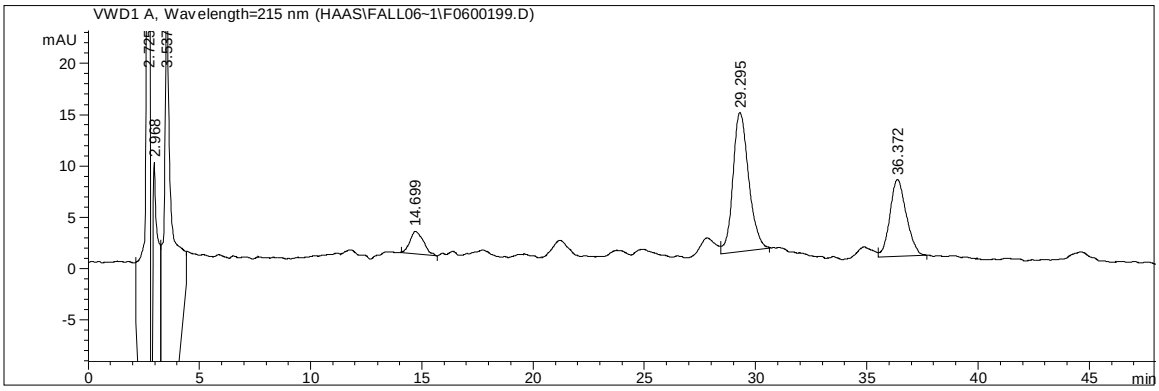
Figure 25. The retention time of *T. maritima* phosphono-CheY nearly overlaps that of *T. maritima* CheY biotinylated with PEO-iodoacetyl biotin and PEO-maleimide biotin but no CheY biotinylated with biotin-HPDP. a.) RP-HPLC of a phosphonomethylation reaction mixture of *T. maritima* D54C/C81S CheY after biotinylation with biotin-HPDP. Phosphono-CheY elutes at 29 min and HPDP-biotinylated CheY elutes at 36 min. b.) RP-HPLC of a PEO-iodoacetyl biotinylation mixture of *T. maritima* CheY. PEO-iodoacteyl biotinylated CheY elutes at 22 min, almost exactly where phosphono-CheY elutes under the same gradient. c.) RP-HPLC of a PEO-maleimide biotinylation of *T. maritima* CheY on a slightly different gradient. Phosphono-CheY elutes at 28-30 min under this gradient. Residual unmodified CheY is seen at 44 min.



a.

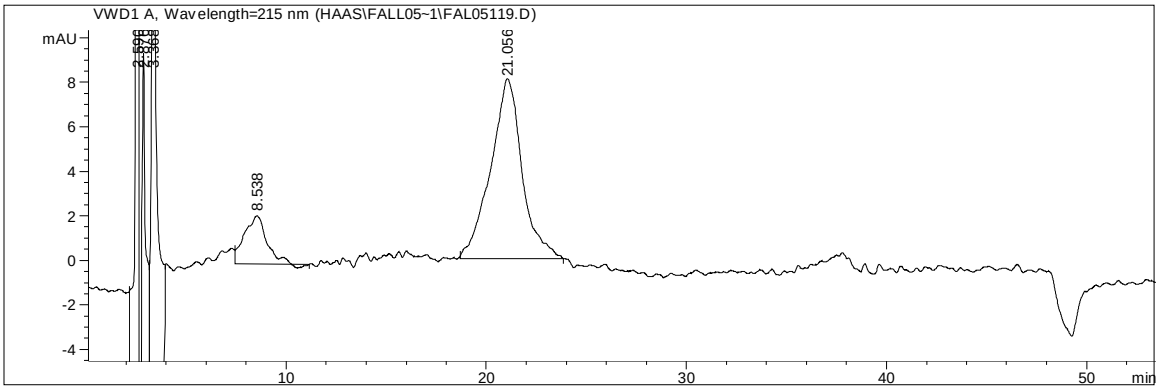


b.

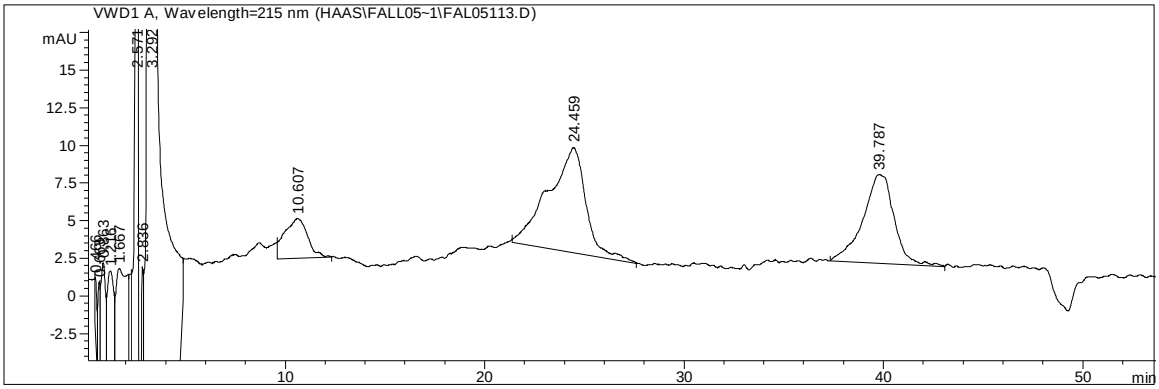


c.

Figure 26. RP-HPLC analysis of the synthesis and purification of *T. maritima* phosphono-CheY sample TMY41'. a.) Single-prime phosphonomethylation reaction mixture. b.) Double-prime sample after biotinylation with biotin-HPDP. c.) Triple-prime sample after the avidin column. Phosphono-CheY elutes at 28 min, biotinylated CheY elutes at 36 min, and unmodified D54C/C81S CheY elutes at 41 min. The biotinylation reaction has gone to 100% completeness in panel b, but a species consistent with biotinylated CheY remains after the avidin column in panel c. See Figure 23c for RP-HPLC of this sample taken after a second pass over avidin.

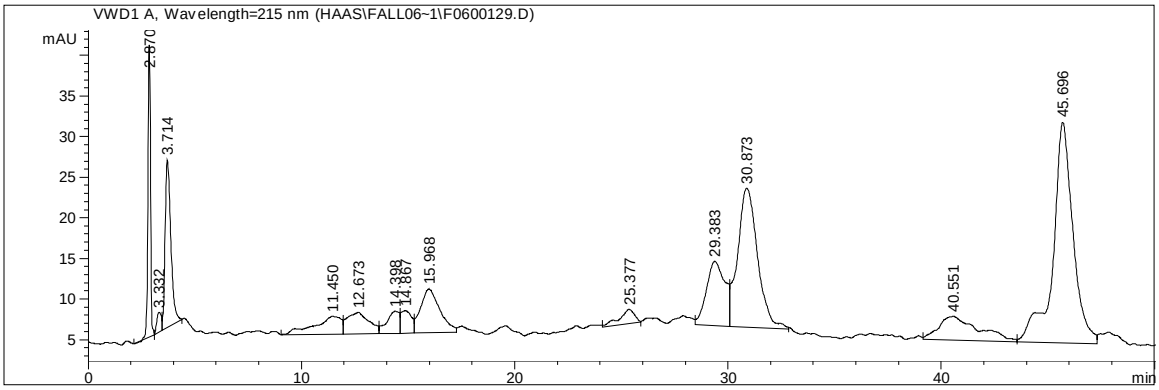


a.

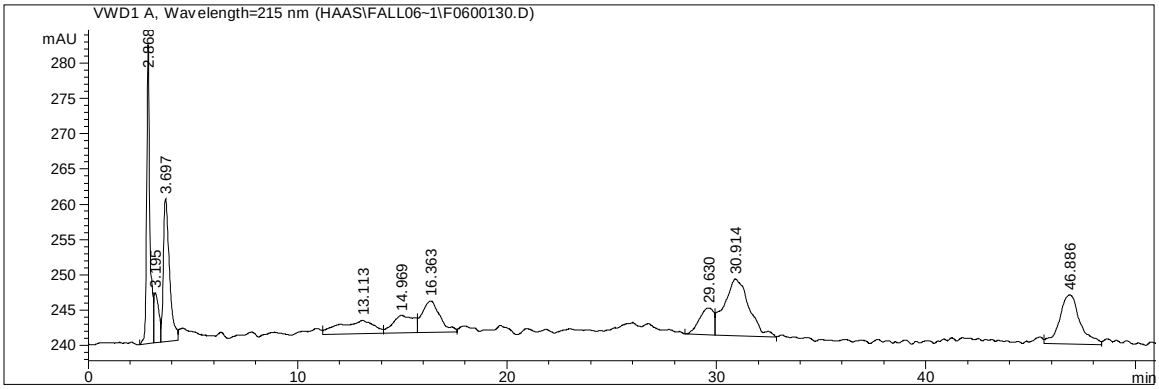


b.

Figure 27. a.) RP-HPLC of the double-prime sample after the biotinylation of a *T. maritima* phosphonomethylation mixture. The peak at 21 min represents both phosphono-CheY and PEO-iodoacetyl-biotinylated CheY. The peak at 8 min is unidentified. Notice the slight hump in the chromatogram near 38 min, suggesting unmodified CheY was not present in the sample. b.) RP-HPLC of this mixture after it was taken over avidin for the first time. A peak at 39 min is consistent with the retention time of unmodified D54C/C81S CheY. This suggests some sort of hydrolysis has cleaved the thioether bond between the protein and the biotin reagent, resulting in the apparent re-emergence of unmodified CheY. This has no chemical precedent, however, and the peak may represent a different species altogether.

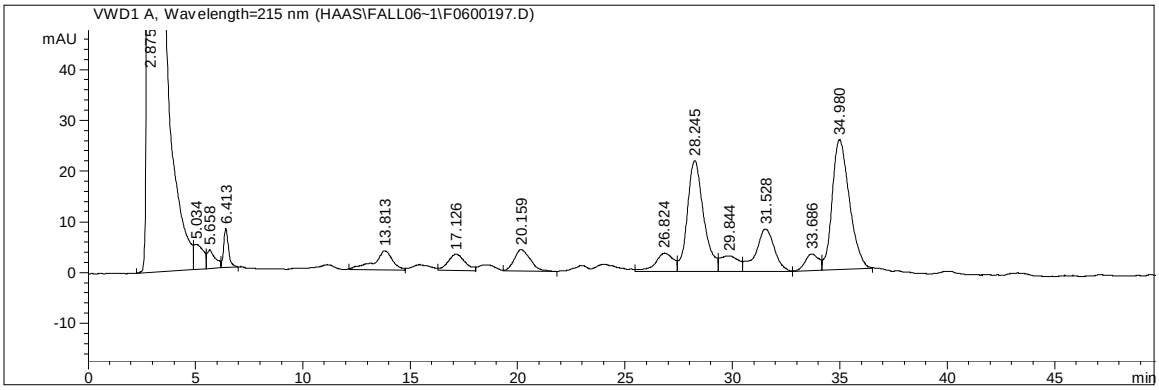


a.

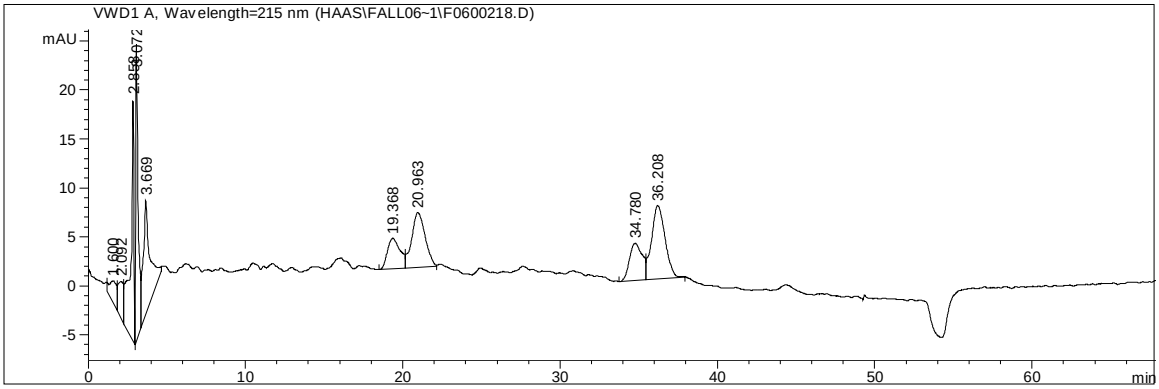


b.

Figure 28. a.) RP-HPLC of a phosphonomethylation mixture of *T. maritima* D54C/C81S CheY. b.) RP-HPLC of the same mixture after biotinylation with PEO-iodoacetyl biotin. In each case, injection consisted of 40 µg of protein in 60 µL. The signal is ~2.5-fold less by area in the analysis of the biotinylation mixture. Phosphono-CheY and iodoacetyl-biotinylated CheY elute at 30 min and unmodified D54C/C81S CheY elutes at 45 min.



a.



b.

Figure 29. a.) RP-HPLC of a *T. maritima* CheY biotinylation mixture (biotin-HPDP). Phosphono-CheY elutes at 28 min and biotinylated CheY elutes at 35 min. b.) RP-HPLC of a different *T. maritima* phosphonomethylation reaction mixture after reacting unmodified CheY with biotin-HPDP. The phosphono-CheY peak which elutes at 28 min is no longer found after biotinylation, and a new species appears as a doublet at 19-20 min.

Purification of *T. maritima* Phosphono-CheY by HPLC

Purification of phosphono-CheY from *Thermotoga maritima* was possible both by reversed-phase HPLC and hydrophobic interaction chromatography. Peaks from reversed-phase were collected in ACN/H_2O/TFA and lyophilized to powder. Reversed-phase conditions may or may not be denaturing to the protein, so phosphono-CheY collected in this manner was used mostly for mass spectrometry experiments and was not shipped to our collaborators. Hydrophobic interaction chromatography, however, is a non-denaturing form of HPLC and phosphono-CheY purified in this manner should be natively folded and ideal for crystallization and binding experiments. In order to purify out phosphono-CheY from the reaction mixture, unmodified CheY was labeled with PEO-iodoacetyl biotin (Figure 30). This shifted the retention time of CheY back approximately 4 minutes and made separation of phosphono-CheY easier since it elutes close to (~0.5 min before) unmodified CheY (Figure 31). The reaction mixture could then be concentrated and mixed with an equal volume of saturated ammonium sulfate and loaded onto the PolyPropyl A column. The column holds between 1 – 5 mg of protein, so 1 – 2 mg of phosphono-CheY (assuming 40% modification) could be purified at a time.

During the HIC purification, everything (1 – 13.5 min) before the CheY peak was collected in one fraction and 1 min fractions were collected after that. Figures 30 and 31 are reversed-phase analyses of the protein present in each HIC fraction. The first 13.5 minutes did not result in phosphono-CheY by reversed-phase analysis. The later fractions contained phosphono-CheY as the dominant species present along with a number of earlier eluting peaks. The odd result is the same relative appearance of all chromatograms, suggesting that the 15.6 and 17.1 min peaks in figure 32 may be phosphono-CheY eluting earlier than it usually does by RP-HPLC. Multiple reversed-phase sample preparations of this fraction, including reduction with 30

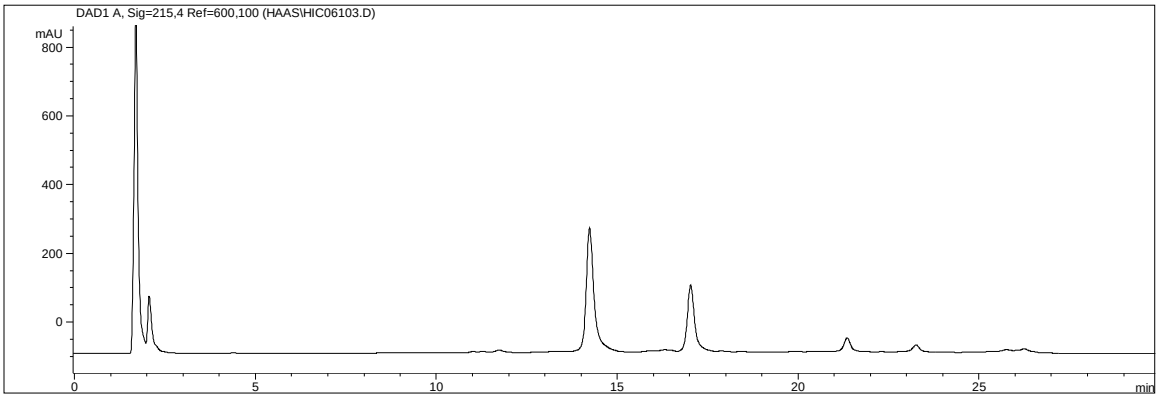
mM DTT, failed to shift the retention time of this species, indicating it is not phosphono-CheY. Moreover, these peaks are consistent both in appearance in approximate retention time with early eluting peaks in some phosphonomethylation reactions (see figure 28a This sample has been submitted for MS analysis. Phosphono-CheY is most likely one of the peaks eluting between 12-14 minutes in figure 31c, and HIC still remains a viable means for purification.

A second non-denaturing HPLC technique was also attempted but was not successful in purifying phosphono-CheY from unmodified D54C/C81S CheY. A weak anion exchange column run under conditions of minimal ionic strength (5 mM Tris-acetate, pH 7.9) was unable to achieve significant binding of either phosphono-CheY or unmodified CheY (Figure 34). If partial purification is achieved after one pass over IEX, it may be possible to apply the sample to the column twice for complete purification. Both of these ideas would require accurate measurement of the amount of time between detection and collection.

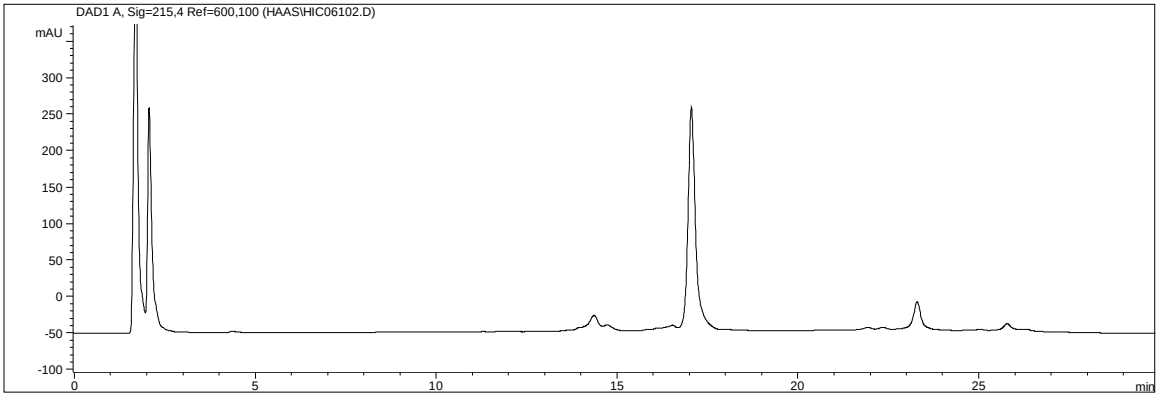
Purification of *T. maritima* Phosphono-CheY with Immobilized Glutathione

Unmodified CheY at 63 µM was reacted with 1.5 mM DTNB (24-fold excess), producing a solution of 76% TNB-labeled CheY as measured by the increase in absorbance at 410 nm. After 24 hours of reaction with immobilized glutathione, the intensity of coomassie-stained bands by SDS-Page indicated 75% of the protein was found in the wash and the remaining 25% was eluted with 2-ME. By Bradford assay, only 400 µg of the 800 µg loaded onto the glutathione column were recovered, 300 µg of which were in the wash. Both samples were run on RP-HPLC and produced a single peak at 44 minutes, the position at which unmodified CheY elutes. The results indicate failure of the TNB-labeled CheY to bind to the glutathione column.

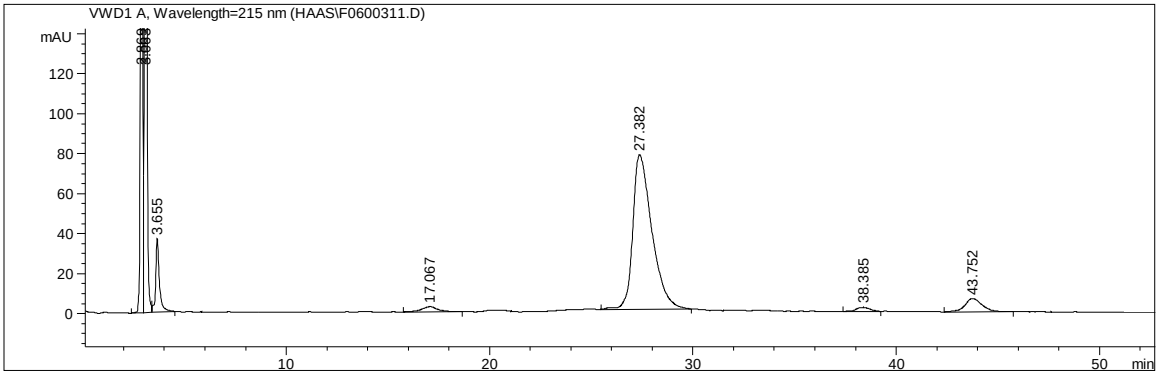
Attempts with a mixture of phosphonomethylated CheY were not made due to the results presented here.



a.

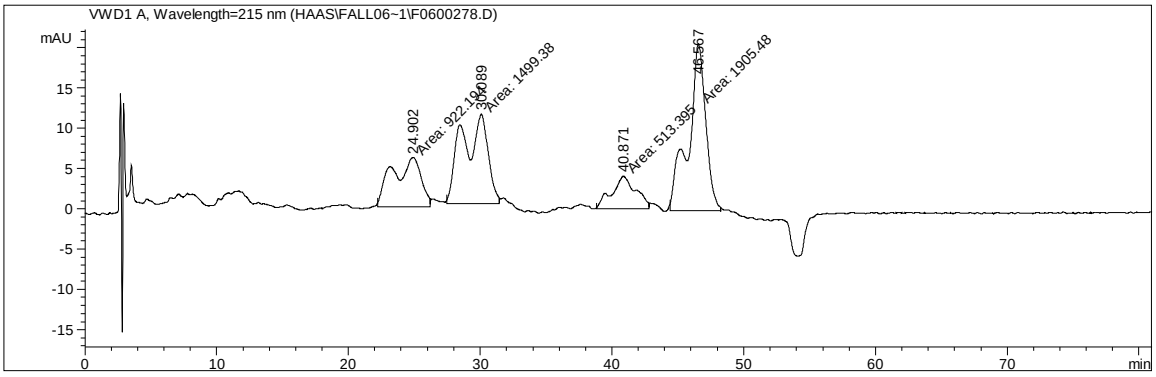


b.

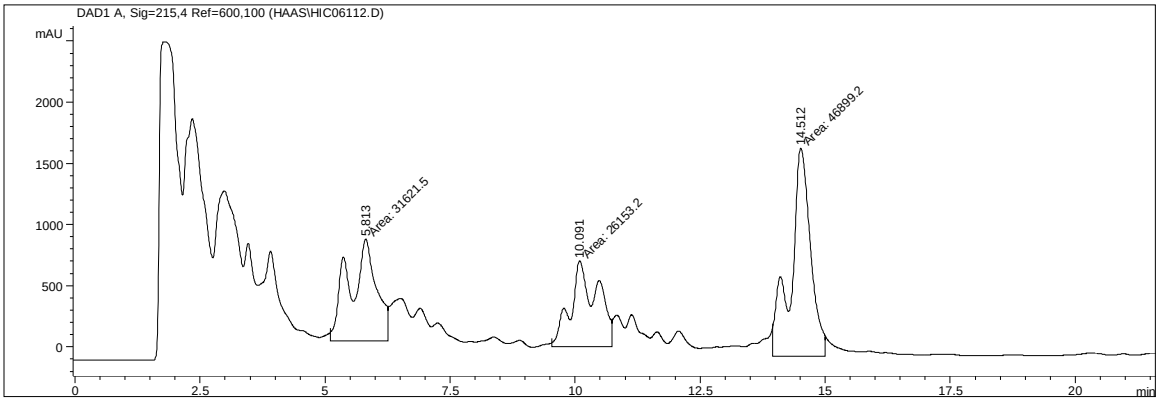


c.

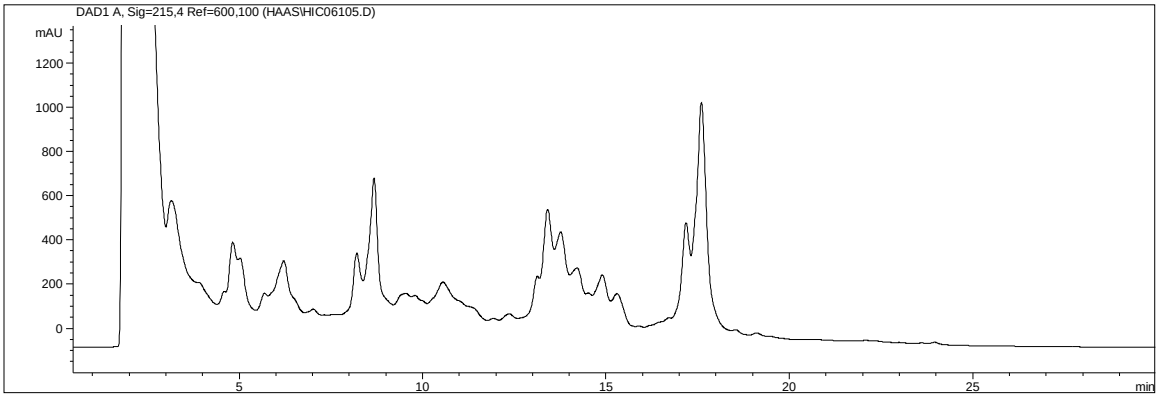
Figure 30. a.) HIC-HPLC of a mixture of *T. maritima* D54C/C81S CheY and PEO-iodoacetyl biotinylated CheY. The 14 min species is unmodified CheY and biotinylated CheY elutes at 17 min. b.) HIC-HPLC of *T. maritima* D54C/C81S CheY biotinylated with PEO-iodoacetyl biotin. The gradient is a decreasing $(NH_4)_2SO_4$ concentration (1.6 M to 160 mM) over 30 min in 20 mM potassium phosphate, pH 6.5. c.) RP-HPLC of the biotinylation mixture depicted by HIC-HPLC in panel b. The biotinylated CheY peak appears at 27 min and the unmodified CheY peak appears at 44 min. The completeness of reaction is consistent between the two HPLC techniques.



a.



b.



c.

Figure 31. a.) RP-HPLC of a phosphonomethylation reaction of *T. maritima* CheY. b.) HIC-HPLC of the same phosphonomethylation reaction mixture. c.) HIC-HPLC of a *T. maritima* biotinylation reaction mixture. The heterogeneity in the phosphonomethylation reaction mixture is evident both by RP-HPLC and HIC-HPLC. The HIC gradient is a decreasing $(NH_4)_2SO_4$ concentration (1.6 M to 160 mM) over 30 min in 20 mM potassium phosphate, pH 6.5. The peak at 14.5 min in panel a is unmodified D54C/C81S CheY and the 17 min peak in panel b is PEO-iodoacetyl biotinylated CheY. It was initially thought that phosphono-CheY was the 14 min peak right before unmodified CheY, but RP-HPLC analysis of the fractions collected suggest phosphono-CheY is present throughout 13.5-17 min. HIC-HPLC of *T. maritima* CheY after phosphonomethylation suggests a great amount of heterogeneity present in the reaction mixture. Buffer components have been ruled out by blank injections and all of the peaks give a characteristic protein spectra.

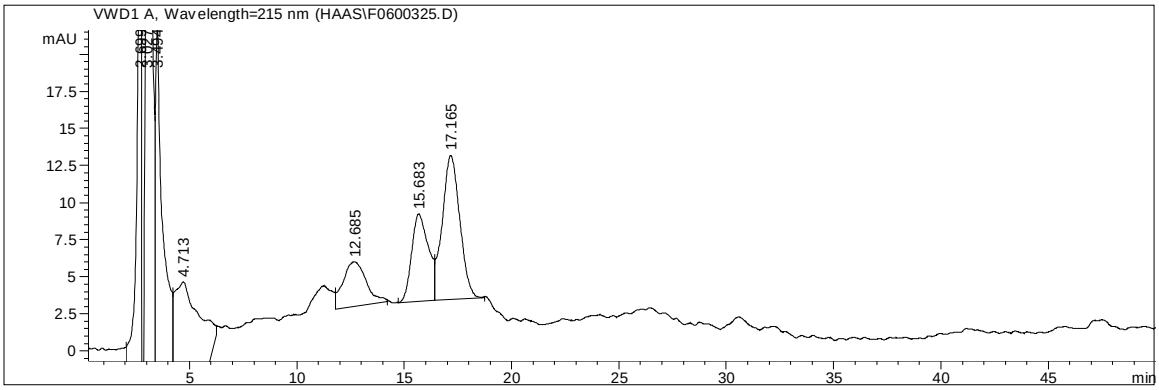
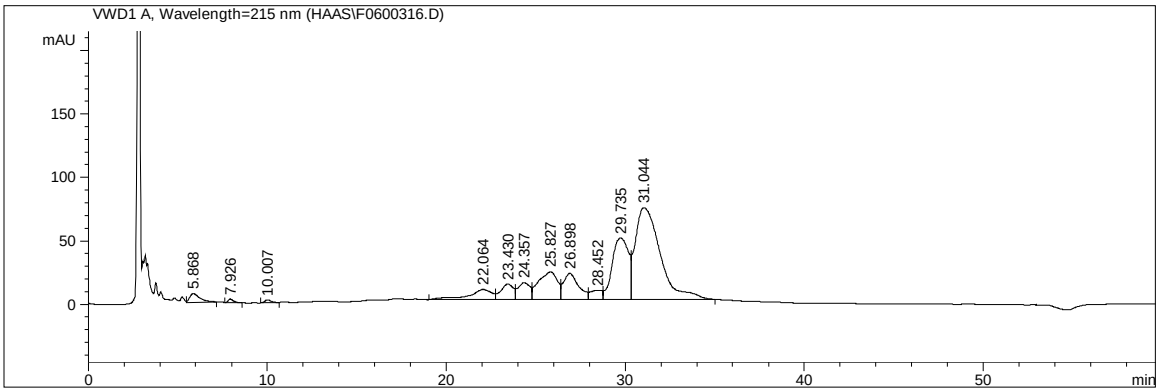
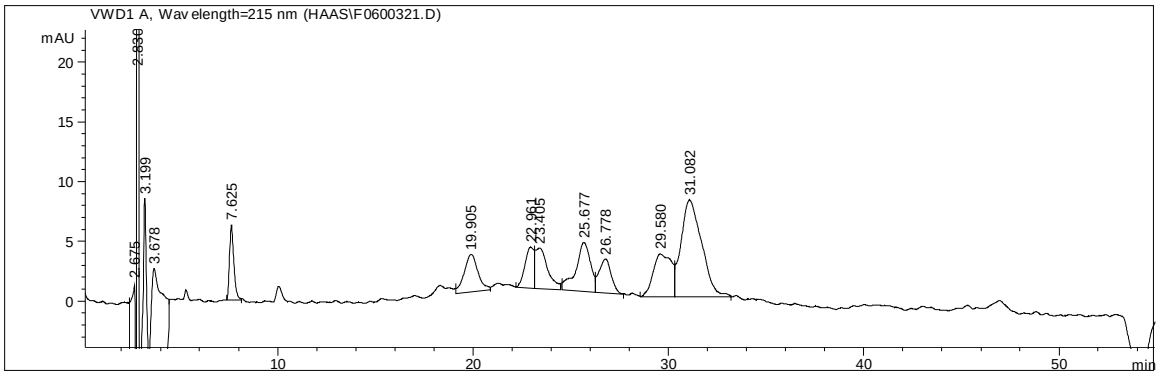


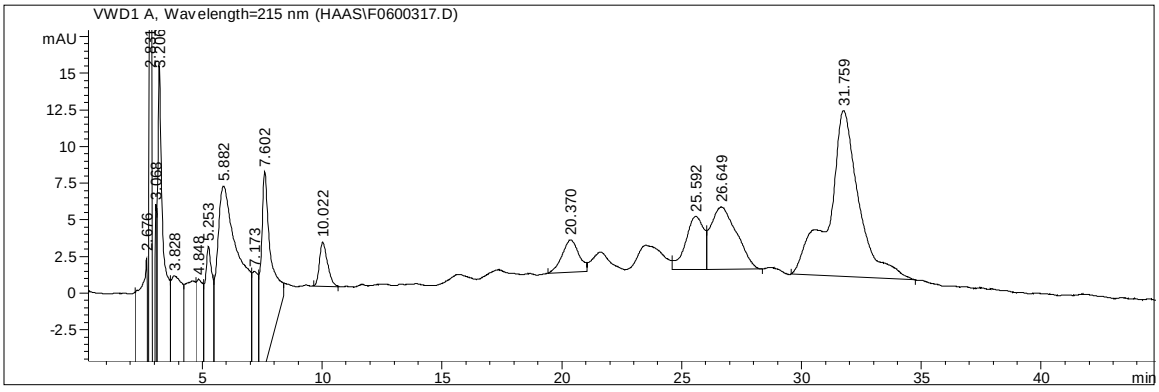
Figure 32. RP-HPLC of the 1-13.5 min fraction collected during the purification of *T. maritima* phosphono-CheY by hydrophobic interaction chromatography (see Figure 31b). Phosphono-CheY elutes at 31 min and is not present in the first 13.5 min of a HIC run. The species present in this chromatogram have not been identified.



a.

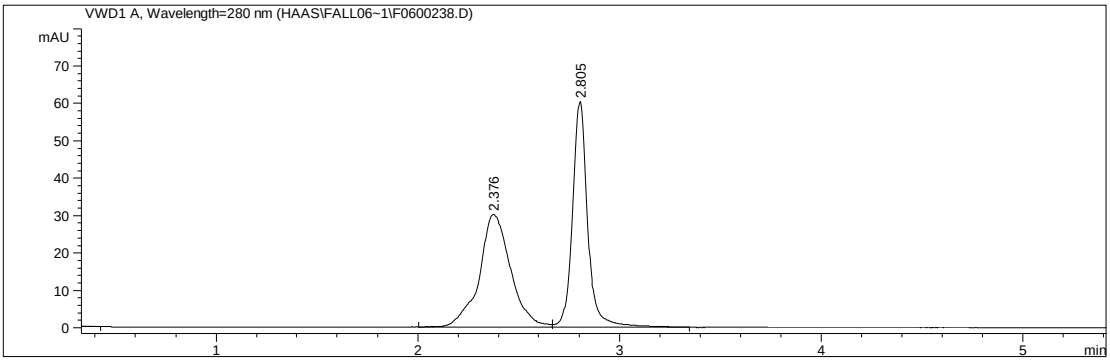


b.

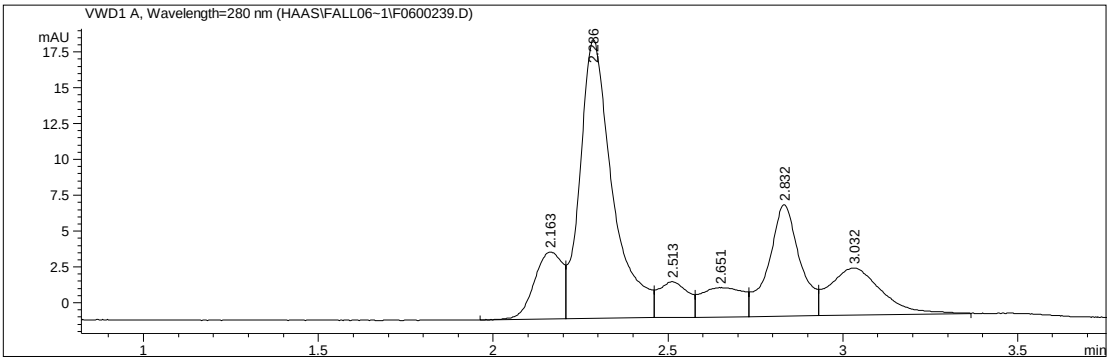


c.

Figure 33. a.) RP-HPLC of the 13.5-15 min fraction collected during the purification of *T. maritima* phosphono-CheY by hydrophobic interaction chromatography (see Figure 31b). b.) RP-HPLC of the 15-16 min fraction collected during the purification of *T. maritima* phosphono-CheY by hydrophobic interaction chromatography. c.) RP-HPLC of the 16-17 min fraction collected during the purification of *T. maritima* phosphono-CheY by hydrophobic interaction chromatography. Phosphono-CheY (31 min peak) is the dominant species present in each of these fractions.



a.



b.

Figure 34. IEX-HPLC of unmodified *T. maritima* D54C/C81S CheY (a) and pure phosphono-CheY (b) illustrate the difficulty in separating the two species from a phosphonomethylation reaction mixture.

Characterization of *T. maritima* Phosphono-CheY by Mass Spectrometry

Intact protein hits an MS detector as multiply charged species and forms a bell-shaped curve known as an "envelope." The mass of the species before ionization can be found by deconvoluting the m/z values according to charge state:

$$m/z = (\text{Molecular Weight} + \text{Charge})/\text{Charge}$$

T. maritima D54C/C81S CheY purified as a single peak by RP-HPLC was found in the +9 to +23 charge states by ESI-TOF MS analysis (Figure 35). Intact *T. maritima* phosphono-CheY collected from RP-HPLC (Figure 36) was found in +10 to +17 charge states by ESI-TOF MS analysis (Figure 37). Unlike the spectrum of unmodified CheY, the spectrum of phosphono-CheY had a low ion count and poor signal-to-noise and as a result contained numerous peaks clustered around each "charge state." Mass reconstruction of the spectrum reveals a number of protein species that are related to CheY either by addition of phosphonomethyl groups (+94) or putative oxidation (+16). This data is the first solid evidence for multiple phosphonomethylations (Figure 38). The oxidized species most likely represent oxidation of methionine to methionine sulfoxide; however, the cascade pattern of decreasing oxidation states is a common phenomenon seen in mass spectrometry of proteins that have seen aqueous buffers. The baseline at 25 min and immediately before phosphono-CheY was also analyzed by mass spectrometry and produced an envelope centered at m/z 978.8 (Figure 39). Mass reconstruction of the peaks in this spectrum provides further evidence for multiple phosphonomethylation of *T. maritima* CheY.

Besides intact protein MS, analysis of peptides was also performed. The peptides generated from tryptic digestion of phosphono-CheY were compared to those from unmodified CheY using gradient LC coupled to ESI-TOF MS. The peptide map generated from the total ion

count of unmodified CheY gave 6 peaks of relatively great intensity and 8 peaks of lower intensity, not including the sharp first peak at 3.25 min (Figure 40). Four peptides were not found in the digest; those with amino acids 38-42, 43-44, 72-77, and 118-120. The peptide of most interest contains residues 48-70, which includes the D54C mutation, and has a mass of 2619.29 Da with the carbamidomethyl-cysteine modification. This peptide was found in only the +3 charge state (m/z 874.0967) with good resolution but poor intensity between 25.1 and 25.5 min. The +2 charge state (m/z 1310.645) was seen with very low resolution at 25.255 min. The +1 charge state (m/z 2620.29), +4 charge state (655.8225), +5 charge states (524.858) were not seen.

The total ion count peptide map of phosphono-CheY is shown in Figure 41 and is not significantly different from unmodified CheY. The same four peptides were absent from this digest as well. The phosphono-peptide has a mass of ~2655 Da and was found with low signal in the +3 charge state (m/z 885.3299) at 23.7 min. Another peptide with the approximate m/z value (886.3043) was also found in +3 charge state between 24.5 and 24.8 min. The data indicate the phosphono-peptide is present, but the poor signal-to-noise suggests it does not ionize very well (or at least doesn't readily form positively charged ions). The presence of the phosphono-peptide would be acceptable as evidence for the phosphonomethyl modification at Cys54 except that the carbamidomethyl-cysteine peptide was also found in the mixture. Moreover, it was found with greater intensity than the phosphono-peptide. This could be the unlikely result of displacement of the phosphonomethyl group by iodoacetamide during sample preparation. It could also be due to sample contamination during the collection of the phosphono-CheY peak from RP-HPLC. A LysC digestion of nearly 100% pure phosphono-

CheY (by RP-HPLC) has been prepared without iodoacetamide alkylation and submitted for MS to rule out these two possibilities.

Mass spectrometry analysis of the tryptic peptides also provided supporting evidence for the presence of multiply phosphonomethylated species. Three peptides were identified in the digest of phosphono-CheY that have an additional 94 mass units and are consistent with phosphonomethylation of two to three different lysine residues. Two peptides, one with m/z 991.9018 and one with m/z 661.5945 represent the +2 and +3 charge state of the peptide containing residues 78-95. A third peptide of m/z 904.9783 represents the +3 charge state of the peptide containing residues 45-67. This provides evidence for three phosphonomethyl modifications—two separate lysine modifications (K117 and either K67 or K47) and the expected cysteine modification (C54)—however, no additional peptide could be found that identifies the site of the fourth phosphonomethyl modification.

Characterization of CheY*

The RP-HPLC impurity peak (CheY*) elutes ~27 min prior to the CheY peak and is present only in UNCW preparations M-T of *T. maritima* D54C/C81S CheY (Figures 11-13). SDS-PAGE indicated that it was a protein of similar molecular weight to *T. maritima* D54C/C81S CheY; however its presence made analytical RP-HPLC of our reaction mixtures significantly more complicated since it was unclear whether it was reactive towards phosphonomethylation. A number of experiments were therefore undertaken in order to characterize this peak and remove it. Ultimately, mass spectrometry provided the most detailed results and suggested CheY dimerization, either by covalent disulfide formation or by non-specific aggregation.

Mass Spectrometry

The individual reversed-phase peak CheY* was analyzed intact by ESI-TOF MS. Two different samples of CheY* were analyzed and two different spectra were obtained (Figures 40 and 41). *T. maritima* CheY is present in both samples, however, each spectra contains a number of other species. Analysis of the spectrum in Figure 42 indicates primarily *T. maritima* D54C/C81S CheY and a cascade of oxidation products. The peaks at 822.7 and 853.7 are unidentified. Analysis of the binomial mass spec data (Figure 43) reveals primarily two species. *T. maritima* CheY is present at mass 13469 Da in the charge states +10 to +21, or in the smaller m/z range of the figure. Looking between the CheY peaks reveals peaks with m/z 929.7, 998.5, 1078.7, 1172.2 Da and so on. These peaks are of significant interest as they deconvolute to a species with an average mass 26936.5 ± 0.5 Da, which is twice the mass of CheY. This tentatively suggests dimerization of CheY, although one cannot make any assumptions as to whether this is a covalent (disulfide) or non-covalent dimer. This is the first evidence for CheY dimerization in our lab, but our collaborators at Cornell have experience with dimerization of other *T. maritima* CheY mutants (Gabriella Gonzales-Bonet, private communication).

Peptide Mapping

The digestion of both homogenous and heterogeneous (with and without the CheY* contaminant) preparations of *T. maritima* D54C/C81S CheY was performed in order to identify any differences in the two protein preparations on the amino acid level (compare figures 13 and 14 for heterogeneous and homogenous protein preparations, respectively). Theoretical digestion with LysC produces 14 peptides, consistent with the 14 peaks found in RP-HPLC analysis (Figure 44). The control experiment allowed three early eluting peaks to be excluded from the peptide count. The chromatogram of the R pool 1 digest (Figure 44a) is identical to that of the

Cornell CheY digest (Figure 44b) with the exception of one late-eluting peak at 33.5 minutes that is larger in the R pool 1 digest. This species occupies 3.7% of the total area in R pool 1 and only 1.1% in Cornell CheY. Similar results were obtained when the experiment was repeated on CheY from R pool 1, M pool 2, and prep D. The peptide map of prep D was similar to that of Cornell CheY (Figure 44c), as was expected from its purity by RP-HPLC (Figure 17b). The peptide map of R pool 1 and M pool 2 were similar to that initially obtained by UNCW CheY.

The CheY* peak and the CheY peak were also digested separately in order to shed light on any difference in their amino acid sequences. Ideally one would obtain the masses of each of these peptides and compare them, but this has not been successful. The peptides generated were qualitatively similar between CheY (Figure 45a) and CheY* (Figure 45b), however, three peaks differ between the samples. A small peak at 21.5 minutes is only in the digestion of CheY* and may be a new peptide. The addition of this peak is consistent with decrease in the area of the 11.9 minute peak and may indicate some sort of covalent modification in this portion of CheY*. Three peaks near 14 min in the CheY digest are found as a single peak in the CheY* digest. If these peptides were terminal peptides, their absence in the digestion of CheY* would indicate a truncated protein. While unlikely, this also could be a lack of resolution in CheY* since two of the 14 minute peaks in the CheY digest add up to approximately the same area as the peak in the CheY* digest. This data suggests that the two species are slightly different at the amino acid level. These samples have been submitted for MS analysis to identify what the differences are.

Reductions

The CheY* peak present in UNCW preparations of *T. maritima* D54C/C81S CheY may be oxidized CheY and a variety of reducing agents were used to disprove this hypothesis. Reduction of O pool 1 with 5 mm TCEP, DTT, BMS, and 22 mM 2-ME failed to reduce the size

of the CheY* peak seen by RP-HPLC. Figure 46a is the un-reduced control experiment after 26 hours. Figures 44b and 44c are after 26 hour reduction with TCEP and DTT, respectively. Results identical to this were obtained with BMS and 2-ME. The peak 5 minutes before the CheY peak is only present in un-reduced protein and is oxidized CheY.

Counter-ion Effect

A counter-ion exchange occurs during RP-HPLC in which the TFA from the mobile phases replaces the buffer salt as the counter-ion to the protein. A variation in hydrophobicity by RP-HPLC could occur within a single-protein solution if incomplete exchange of the counter-ion were to occur. To test this effect, heterogeneous protein was exchanged into the three different gel filtration buffers used in the final step of purification of *T. maritima* D54C/C81S CheY and the samples were prepared with 0.1% and 0.4% TFA for injection on RP-HPLC. Varying the percentage of TFA in HPLC samples had only negligible effects on the appearance of the CheY* peak. Samples prepared with an increased amount of TFA (0.4%) had overall decreased peak widths and slightly lower area percentage of the CheY* peak (Table 5). Interestingly, the area percentage of the CheY* peak decreased as it was prepared in 50 mM Tris, phosphate, and MOPS buffers, respectively. The CheY* peak was the widest and had the least area when the sample was prepared in phosphate buffer with 0.4% TFA, however, the decrease in area may be an artifact of the integration. The counter-ion effect did not appear to the cause of the two populations seen in RP-HPLC.

His-tagged Protein

There was a slight possibility that the protein species eluting early in RP-HPLC was his-tagged CheY despite the presence of a single band by SDS-PAGE after thrombin digestion. In order to test this theory, a thrombin digestion was followed to completion by RP-HPLC and

SDS-PAGE. Two peaks elute in the RP-HPLC shallow gradient of hi-tagged O pool 2 protein (Figure 47a). Digestion of this protein produced two additional species that elute 2.0 and 3.2 minutes later than the first and second peaks, respectively (Figure 47b), and most likely represent CheY and CheY* without the Histidine tag. Three hours after thrombin addition, the two new species are found roughly in the same proportion as their parent species in the time 0 sample; however, this is not the case with the 5 hour and 23 hour samples. The size of the first peak shows a considerable decrease (compared to the late-eluting species) 5 hours after thrombin digestion (Figure 48a) but appears to increase by 23 hours (Figure 48b). In addition, an unidentified peak appears immediately after the late-eluting species in the 5 hour sample. Digestion was followed in parallel by SDS-Page and had essentially reached completion by 5 hours. No new protein species other than the his-tag (2.1 KDa) and CheY (13.5 KDa) were present by coomassie staining. CheY* was not residual his-tagged CheY.

Separation by Anion Exchange

T. maritima D54C/C81S CheY has a basic isoelectric point near pH 8.7 and should remain neutral near that pH. An anion exchange column was run at pH 8.6 in an attempt to remove the CheY* from CheY. The column would separate the two species only if CheY* had a lower pI and remained negatively charged under the chromatographic conditions. CheY* must have an isoelectric point similar to CheY since neither of the protein species stuck to the column at this pH. Fractions 2 and 3 from the low salt wash of the DE-52 anion exchange column combined for 3.1 mg of protein. The load and fractions 1 and 4 gave poor looking spectra with no significant absorbance at 280 nm and were not examined. The 2 M NaCl rinse did not appear to elute any protein, since spectra of fractions 1-3 were poor, and the absorbance of fraction 4 slowly descended as it was scanned from lower wavelength to higher wavelength. However, the

2 M salt elute did contain protein, but to a much lesser extent than the low salt fractions. Less than 1.1 mg of protein was lost over the column in this experiment. Both the CheY peak and the CheY* peak appeared upon RP-HPLC analysis of low salt fractions 2 and 3 (Figure 49a). The CheY* was slightly greater in area than the CheY peak (51% vs. 49%). Also a slight shift of 1.5 min to an earlier retention time was observed with this fraction. A very small peak, barely above baseline, appeared at the 20-21 minute mark, where peak 1 normally elutes. This is most likely of no significance. Both peaks were present in RP-HPLC analysis of the 2 M NaCl fractions (Figure 49b). The area percent of the CheY peak decreased by 5% relative to what was observed in the low salt fractions. Interestingly, the percent area of the CheY peak in both cases was lower than it was initially (Figure 14b). Peak 1 represented 38% of the total area and peak 2 was 62%.

Quantitation of *T. maritima* D54C/C81S CheY

The A_{205}/A_{280} method and the Bradford assay gave almost identical results for the concentration of the unknown *T. maritima* sample. The BCA assay, which is normally a more accurate colorimetric assay than the Bradford assay, overestimated the concentration of the *T. maritima* sample by 1.3-fold compared to the other two methods. The calculations associated with the A_{205}/A_{280} method are described below. The absorbances at 280 nm and 205 nm were obtained on the Cary spectrophotometer and are listed after dilution corrections for BSA, egg albumin, and *T. maritima* phosphono-CheY. From these absorbances, the specific absorbance coefficient (or the absorbance of a 1 mg/mL solution at 205 nm) $A_{205}^{0.1\%}$ can be calculated. The literature value for BSA is 29.7 (*52*).

BSA

$$A_{280} = 0.3338 \text{ x } \left(\frac{1.000mL}{0.500mL}\right) = 0.6676$$

$$A_{205} = 0.4398 \text{ x } \left(\frac{1.015mL}{0.015mL}\right) = 29.7598$$

$$A_{205}^{0.1\%} = 27 + 120\left(\frac{A_{280}}{A_{205}}\right) = 27 + 120\left(\frac{0.6676}{29.7598}\right) = 29.69$$

Egg Albumin

$$A_{280} = 0.3671 \text{ x } \left(\frac{1.000mL}{0.500mL}\right) = 0.7342$$

$$A_{205} = 0.4256 \text{ x } \left(\frac{1.015mL}{0.015mL}\right) = 28.80$$

$$A_{205}^{0.1\%} = 27 + 120\left(\frac{A_{280}}{A_{205}}\right) = 27 + 120\left(\frac{0.7342}{28.7989}\right) = 30.06$$

TMY 2'

$$A_{280} = 0.269$$

$$A_{205} = 0.390 \text{ x } \left(\frac{1.010mL}{0.010mL}\right) = 39.39$$

$$A_{205}^{0.1\%} = 27.82$$

One is then able to estimate the protein concentration based on Beer's Law, $A = \varepsilon bc$. The stock solutions of BSA and egg albumin were 1.0 mg/mL. The concentration of Tmy 2' was unknown.

BSA $c = A_{205}/\varepsilon_{205} = 29.76/29.69 = 1.0$ mg/mL

Egg Albumin $c = A_{205}/\varepsilon_{205} = 28.80/30.06 = 0.96$ mg/mL

TMY 2' $c = A_{205}/\varepsilon_{205} = 39.39/27.82 = 1.42$ mg/mL

In addition, one can use the following equation to calculate the specific absorbance coefficient at 280 nm (*53*):

$$A_{280}^{0.1\%} = (27 \times A_{280}/A_{205}) / (1 - 3.85 \times A_{280}/A_{205})$$

$A_{280}^{0.1\%}$ is the absorbance of a 1 mg/mL protein solution at 280 nm. See Table 6 for calculated molar absorptivities at 280 nm. The literature value and its reference are listed within parentheses next to the calculated value.

The Better Bradford assay was also used to determine the concentration of TMY 2'. 10 uL of the protein in 2.1 ml total volume gave an $A_{595} = 0.611$. Using the equation derived from the standard curve: $A_{595} = 0.0303 * C + 0.4086$, where C is concentration in µg/mL.

(0.611-0.4086)/0.0303 = *6.68 ug/mL* * 2.1 mL = *14.028* µg/10µL = 1.40 mg/mL.

The BCA assay was used in an identical manner to the Bradford Assay to determine the concentration of TMY 2'. 10 uL of the protein in 2.1 ml total volume gave an $A_{560} = 0.832$. Using the equation derived from the standard curve: $A_{560} = 0.0435 * C + 0.4531$, where C is concentration in µg/mL.

(0.781-0.4531)/0.0435 = 8.71 µg/mL * 2.1 mL = 18.29 µg/10 uL = 1.83 mg/mL

Results from the Better Bradford assay were consistent with the data obtained by the A_{205}/A_{280} method. The BCA assay overestimated the protein concentration by 1.3-fold relative to the other methods used.

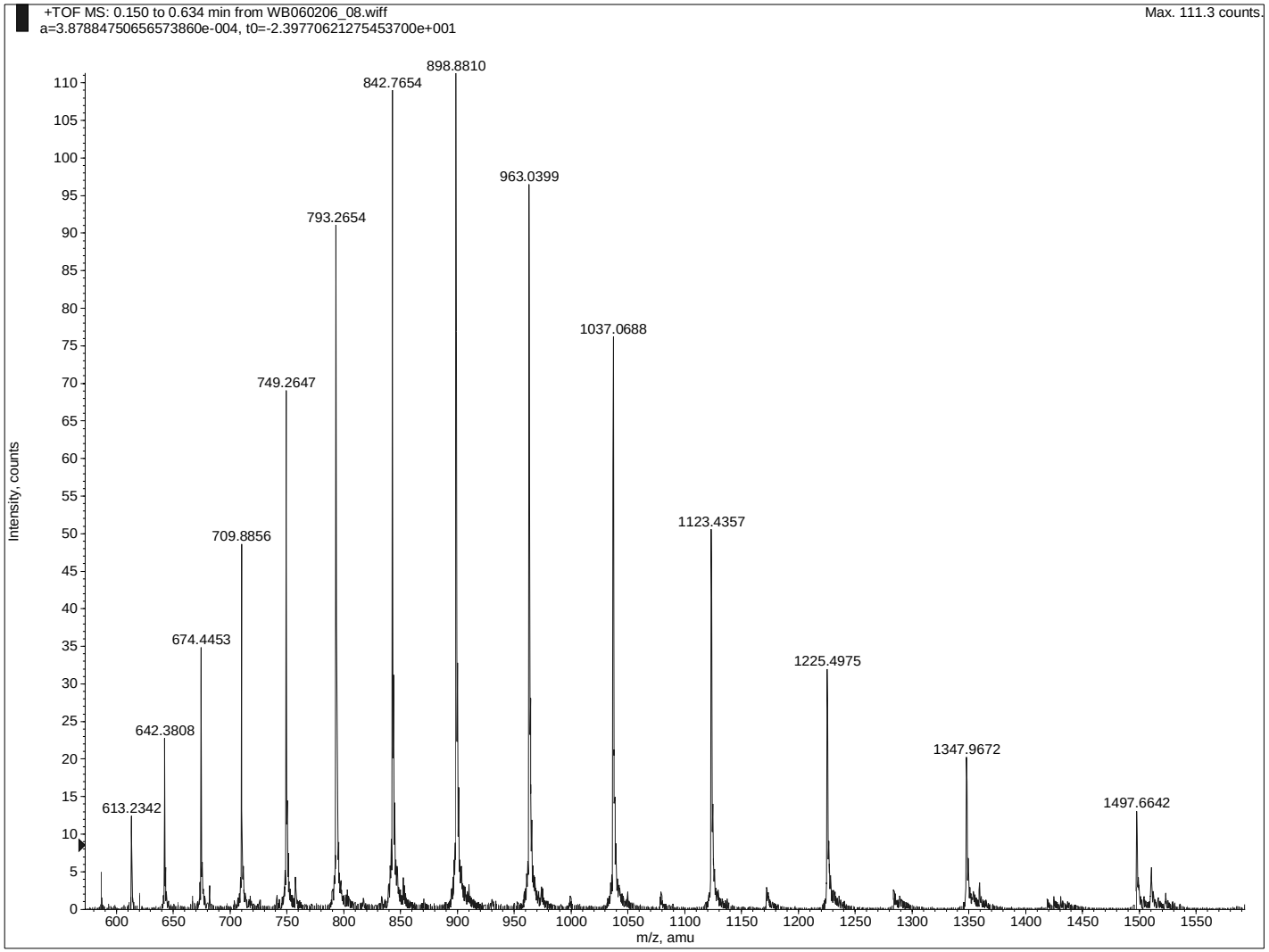


Figure 35. ESI-TOF MS analysis of pure *T. maritima* D54C/C81S CheY. From right to left are the m/z values of the +9 to +22 charge states.

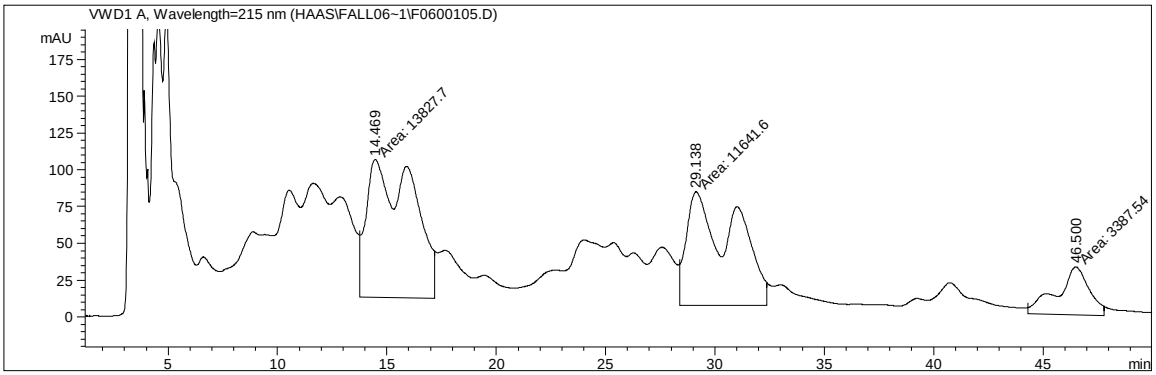


Figure 36. RP-HPLC of a *T. maritima* phosphonomethylation reaction mixture. Phosphono-CheY was collected from 29-33 min for intact mass spec analysis and tryptic digestion. The baseline is not flat before the 29 min peak and most likely led to the presence of species other than phosphono-CheY in the mixture. The eluate before the 29 min peak was also collected for mass spec analysis and contained species also present in the 29 min peak.

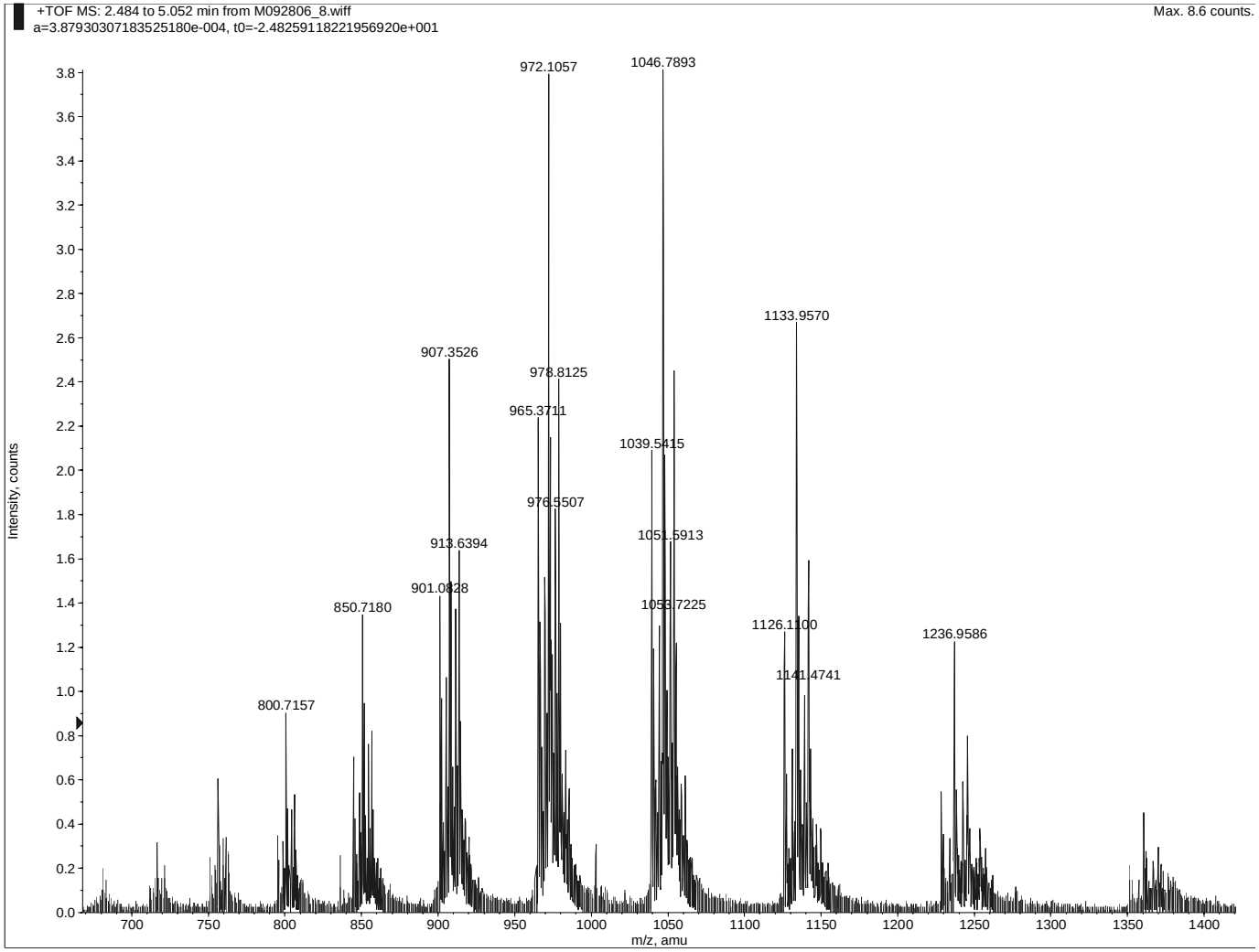


Figure 37. ESI-TOF MS Analysis of Full-Length *T. maritima* Phosphono-CheY. An envelope containing the +10 to +17 charge states of the protein is found.

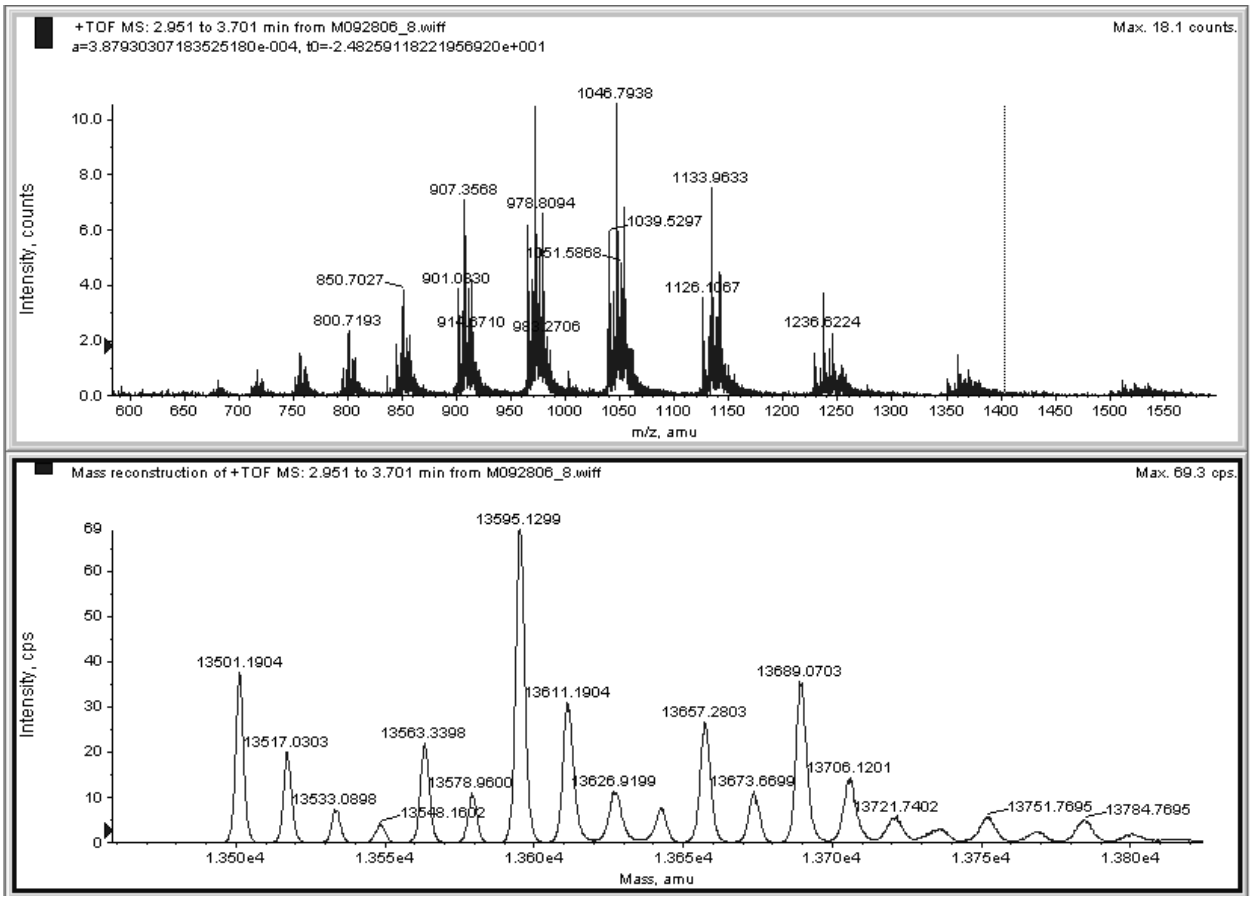


Figure 38. Deconvoluted ESI-TOF MS of the 29 minute presumed phosphono-CheY peak in Figure 40. This spectrum is also illustrated without mass reconstruction in Figure 36. Phosphono-CheY has a mass of 13563 Da. The species with masses 13657 and 13751 represent two and three phosphonomethyl additions (+94 Da) to unmodified *T. maritima* D54C/C81S CheY (mass 13469 Da). The species at 13501 Da is the addition of 32 Da to unmodified *T. maritima* D54C/C81S CheY, possibly indicated oxidation of methionine.

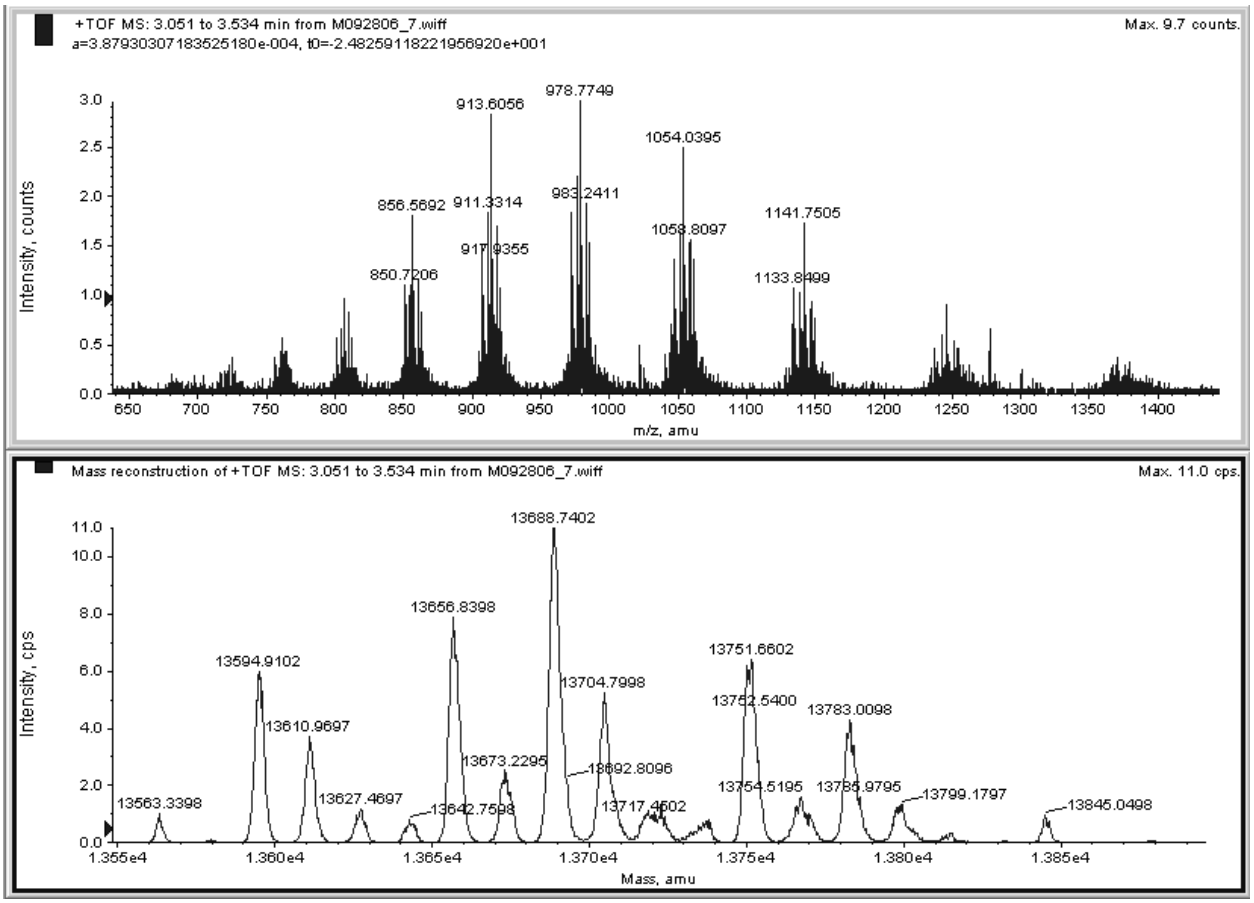


Figure 39: Deconvoluted ESI-TOF MS of the eluate before the 29 minute presumed phosphono-CheY peak in Figure 40. Phosphono-CheY has a mass of 13563 Da. The species with masses 13657, 13751, and 13845 Da represent two, three, and four phosphonomethyl additions (+94 Da), respectively, to unmodified *T. maritima* D54C/C81S CheY (mass 13469 Da).

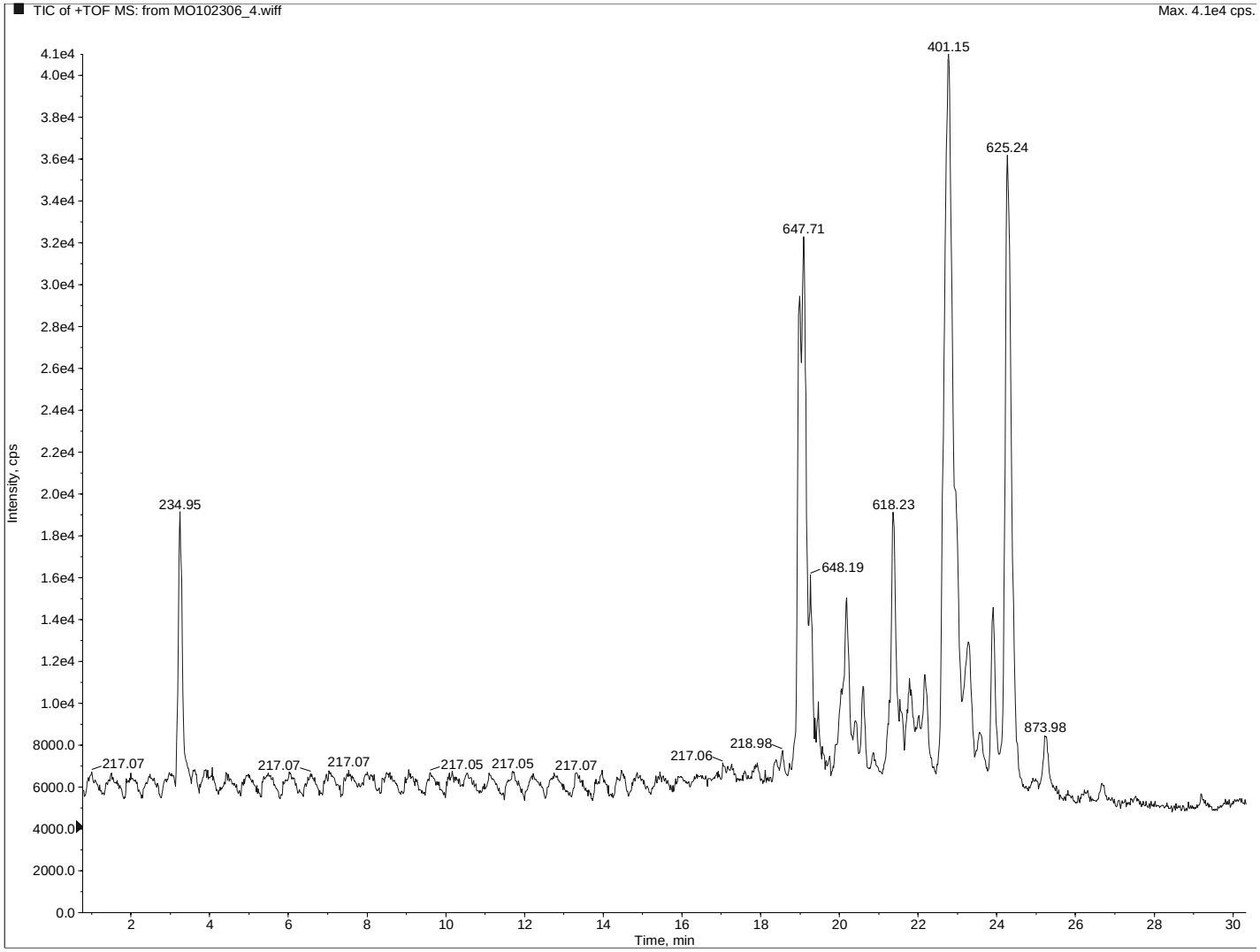


Figure 40. Total ion count peptide map generated by ESI-TOF MS analysis of the tryptic peptides of unmodified *T. maritima* D54C/C81S CheY.

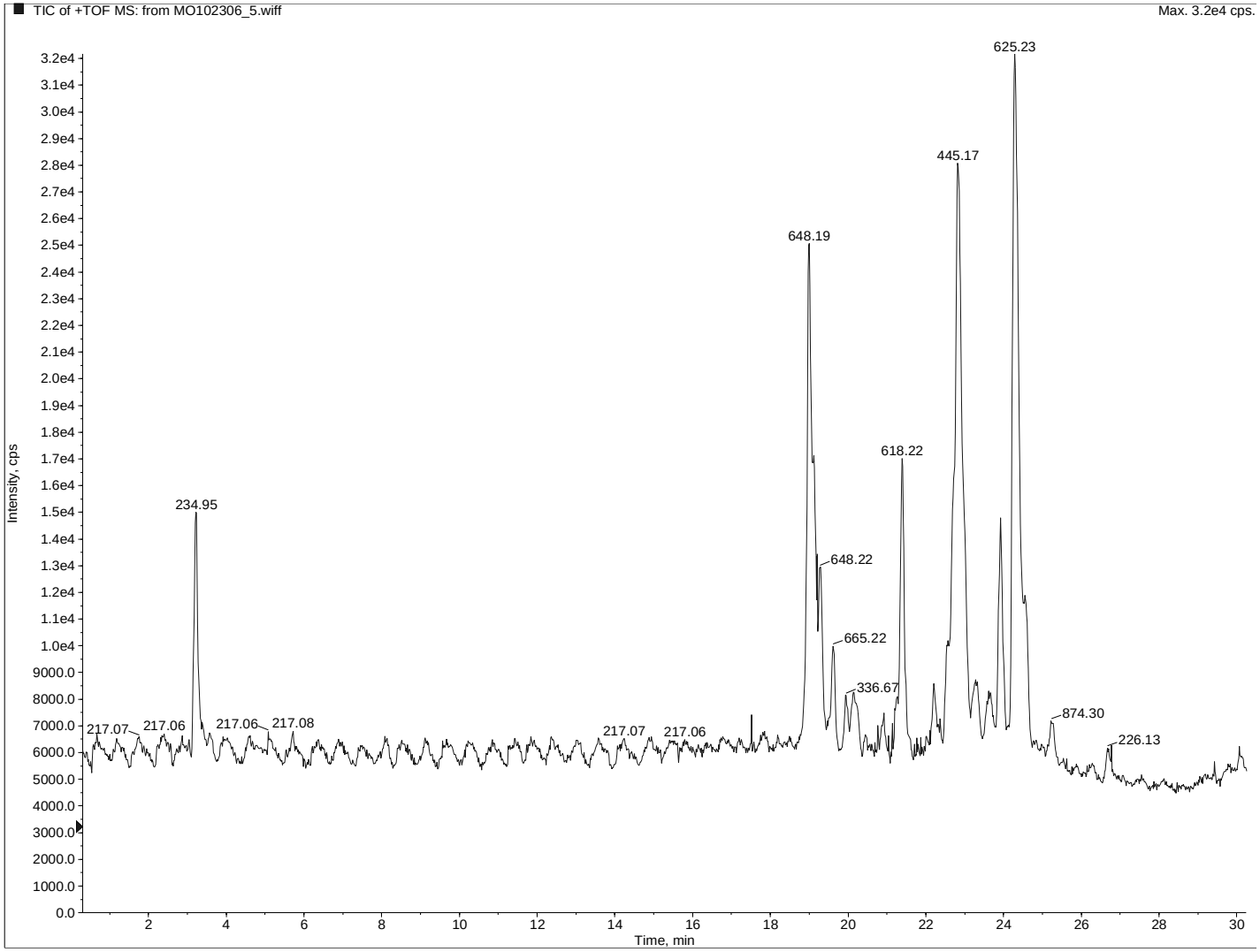


Figure 41. Total Ion Count Peptide Map generated by ESI-TOF MS analysis of the tryptic peptides of *T. maritima* Phosphono-CheY.

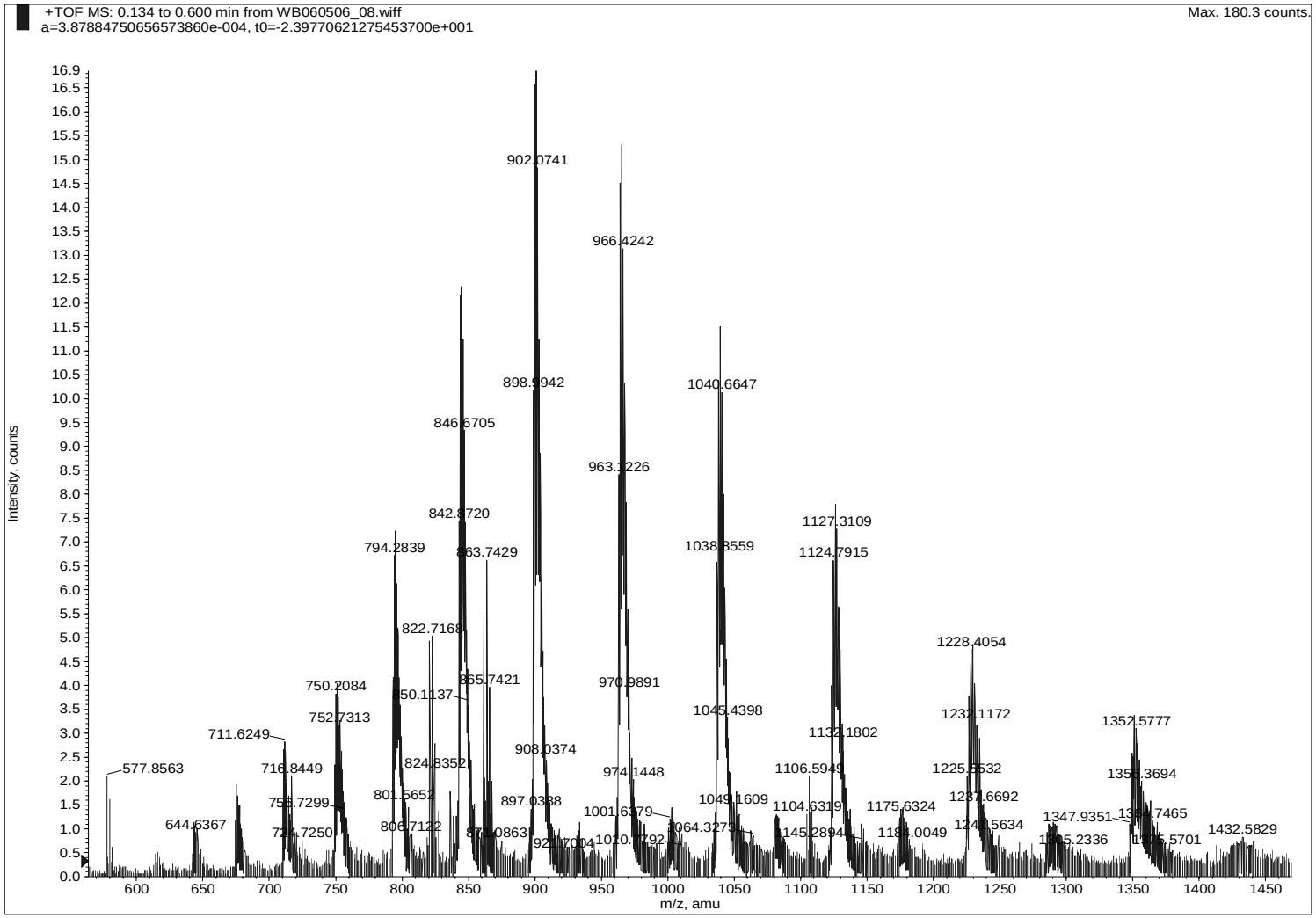


Figure 42. ESI-TOF MS analysis of CheY*.

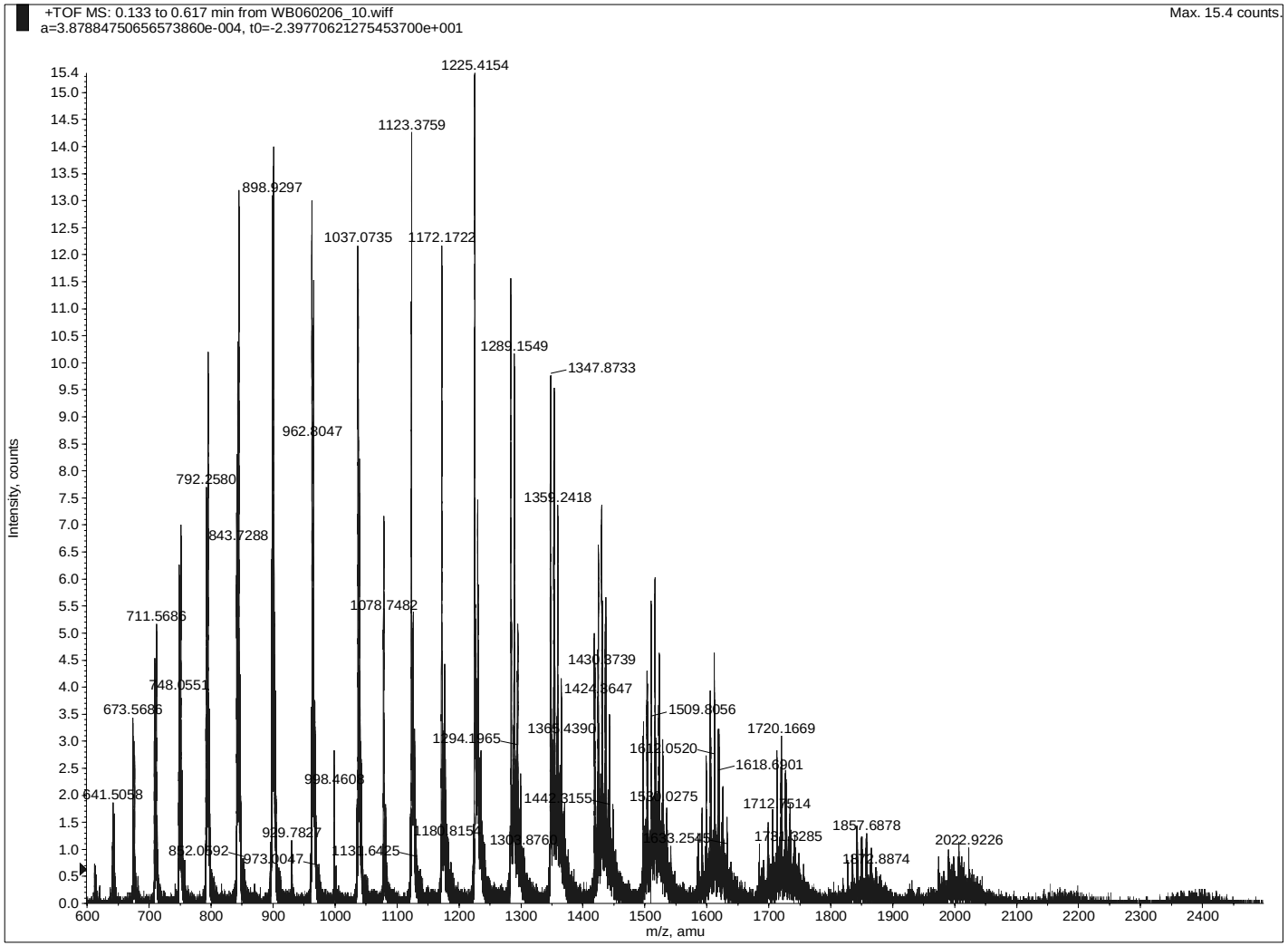
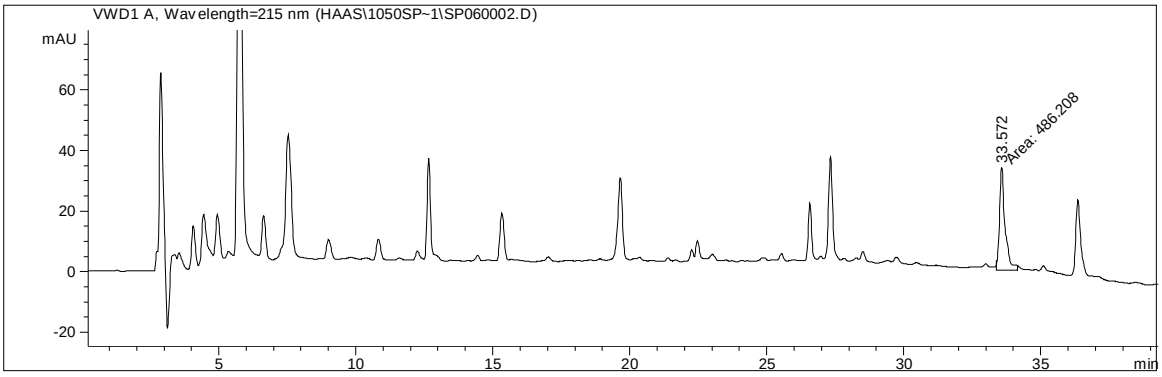
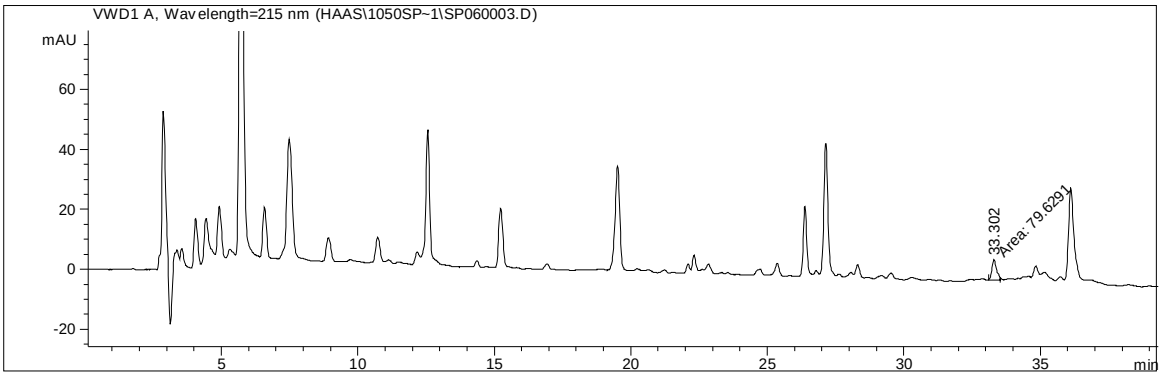


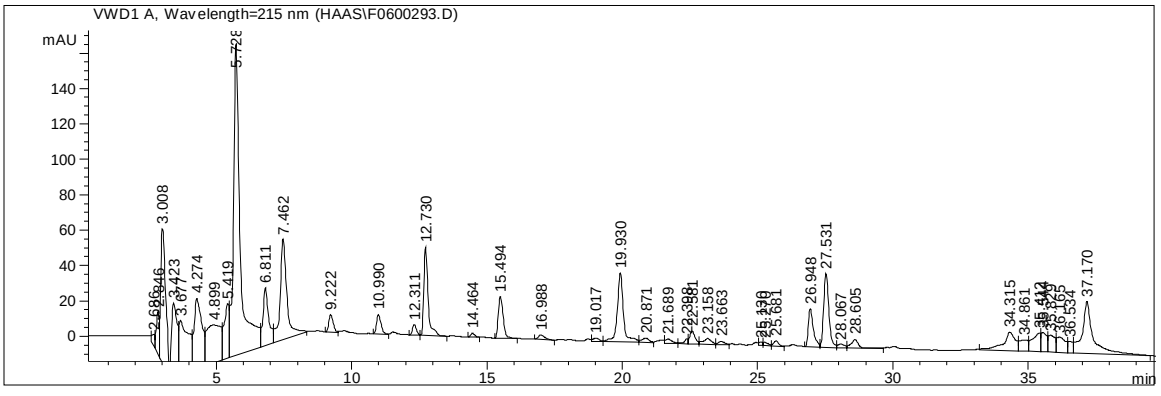
Figure 43. ESI-TOF MS analysis of CheY*. (*T. maritima* D54C/C81S CheY) is present in the +10 to +21 charge states. This spectrum suggests CheY* may be CheY dimerization but gives no information as to whether it is a covalent or non-covalent dimerization.



a.

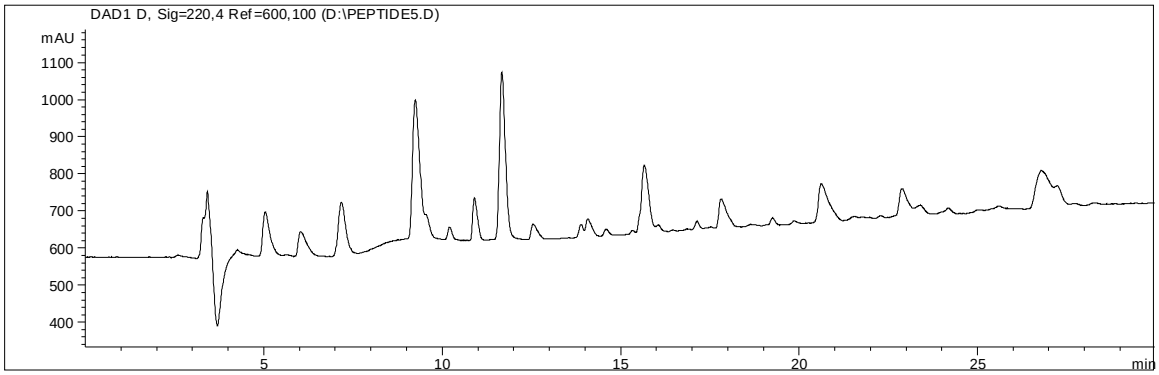


b.

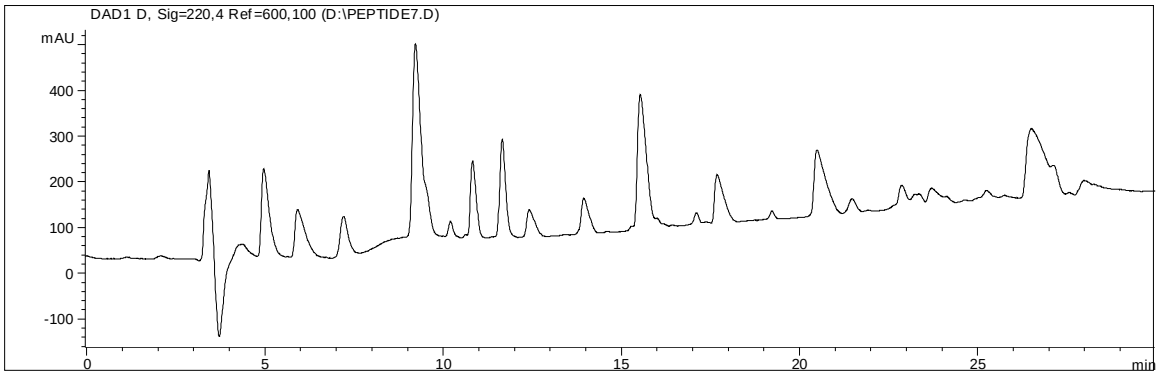


c.

Figure 44. a.) RP-HPLC of the LysC digestion of R pool 1 (heterogeneous *T. maritima* D54C/C81S CheY). b.) RP-HPLC of the LysC digestion of Cornell CheY (homogeneous *T. maritima* D54C/C81S CheY). c.) RP-HPLC of the LysC digestion of *T. maritima* CheY prep D (see Figure 17b for Prep D on a shallow RP-HPLC gradient). Digestion of the two samples produced 14 peaks in both digests. The 33.5 min peak is larger in the digestion of the UNCW preparation of *T. maritima* CheY than in Cornell-prepared CheY.



a.



b.

Figure 45. a.) RP-HPLC of the LysC digestion of CheY from M pool 2 (see Figure 13a for Prep M pool 2 on a shallow RP-HPLC gradient). b.) RP-HPLC of the LysC digestion of CheY* from M pool 2. The gradient was a 30 min, 0-60% mobile phase B where mobile phase B was 95% ACN, 0.1% FA, 0.01% TFA in water and mobile phase A was 5% ACN, 0.1% FA, 0.01% TFA in water.

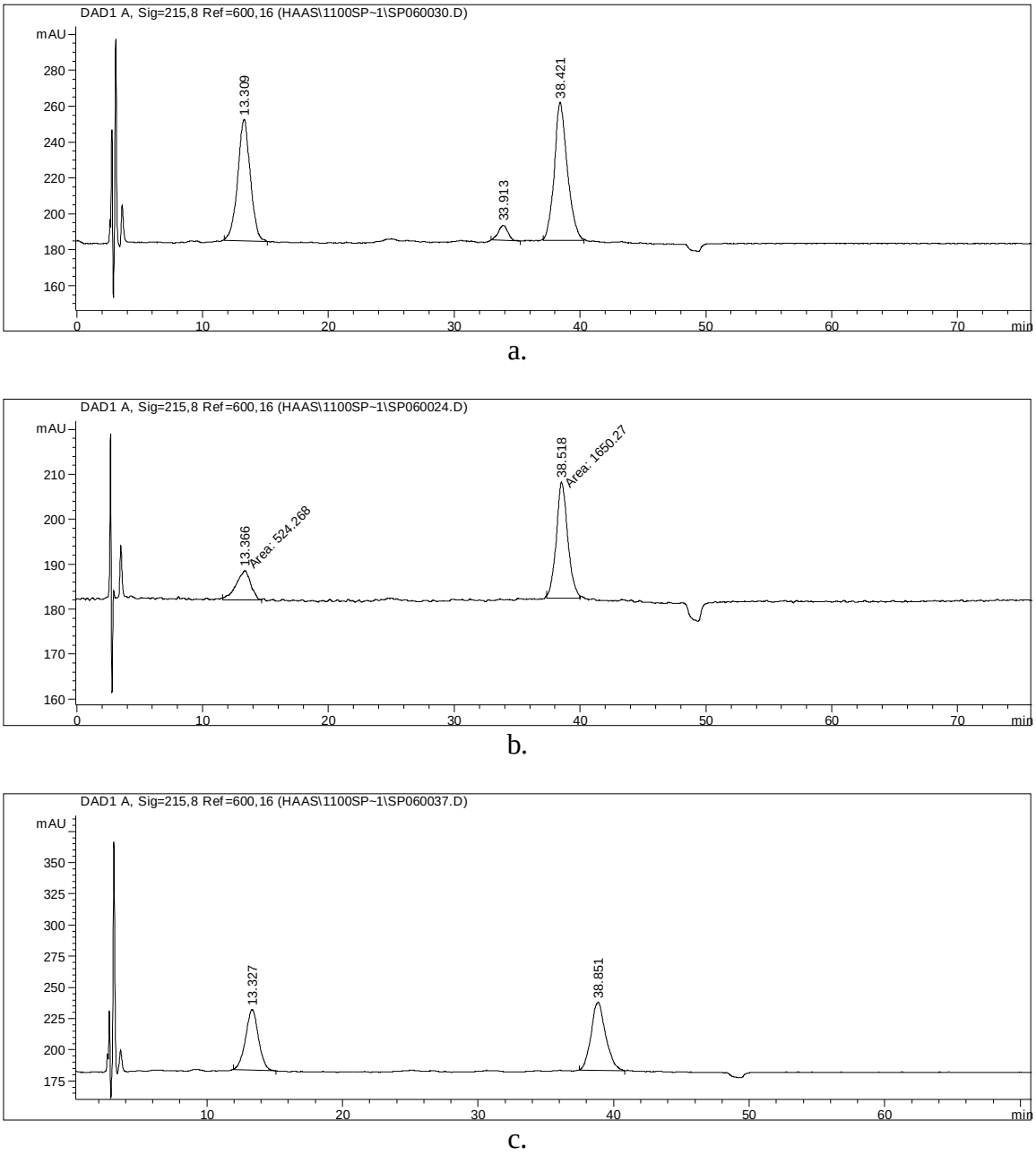
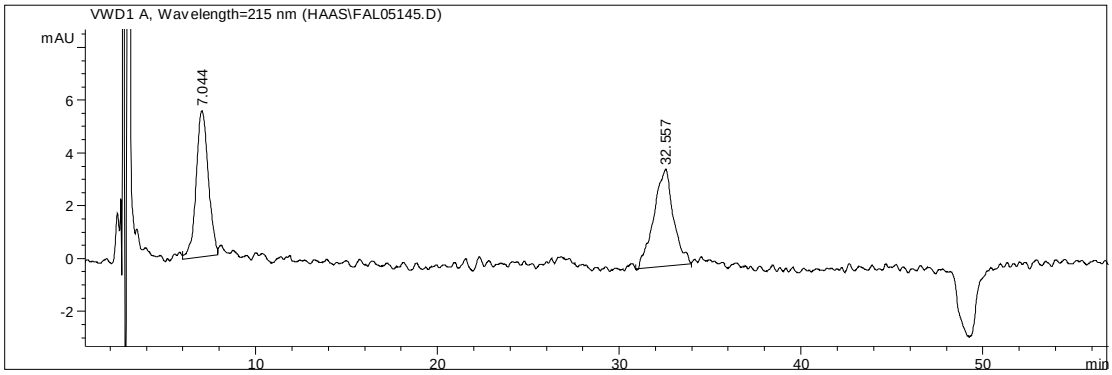


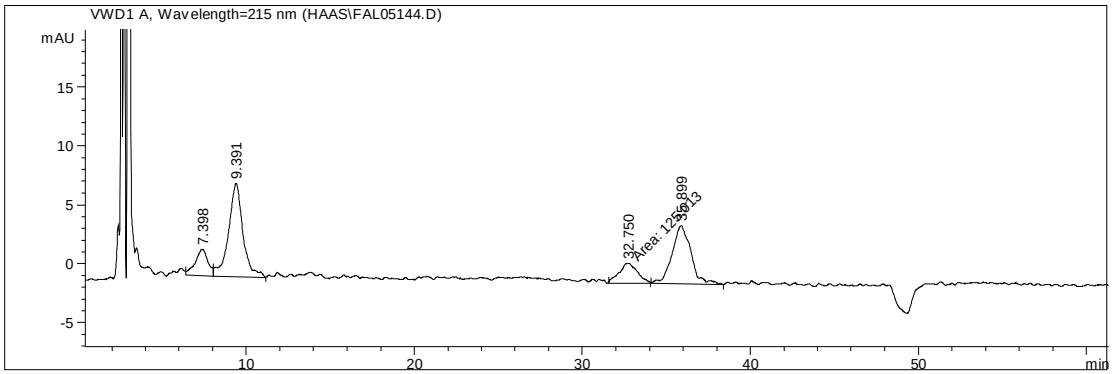
Figure 46. a.) RP-HPLC of the un-reduced control experiment of *T. maritima* D54C/C81S CheY from preparation O pool 1 after 26 hours. b.) RP-HPLC of *T. maritima* D54C/C81S CheY from preparation O pool 1 after 26 hours of TCEP Reduction. c.) RP-HPLC of *T. maritima* D54C/C81S CheY from preparation O pool 1 after 26 hours of DTT Reduction. The CheY* peak at 13 min and the CheY peak at 38 min remain throughout the reductions. TCEP appears to decrease the area of the CheY*. Putatively oxidized CheY appears at 34 min only in the un-reduced control.

Buffer and %TFA	Width/Height	Area% of CheY* Peak
Mops 0.1	0.10117	47
Tris 0.1	0.03307	49
Phosphate 0.1	0.07711	45
Mops 0.4	0.01765	42
Tris 0.4	0.03137	47
Phosphate 0.4	0.30725	38

Table 5. The Counter-ion effect. Increasing the amount of TFA in the HPLC sample decreased the area of the CheY* peak only slightly, but demonstrated that this peak was not the result of incomplete counter-ion exchange during RP-HPLC runs.

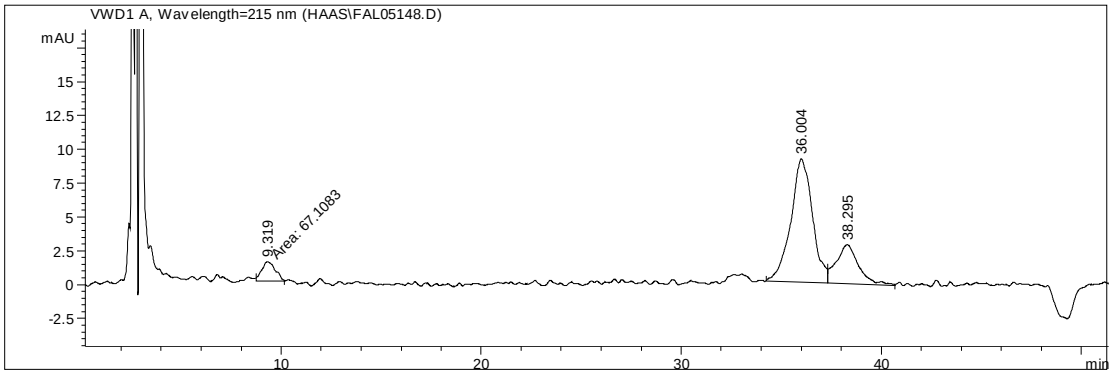


a.

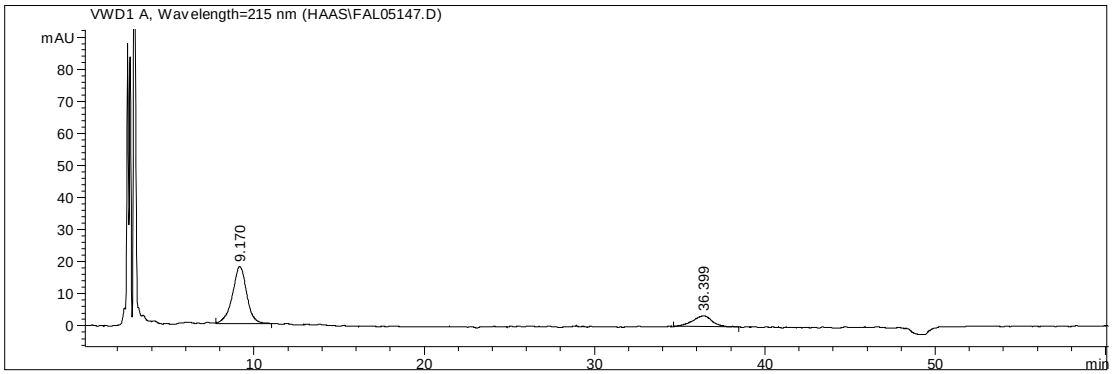


b.

Figure 47. a.) RP-HPLC of his-tagged *T. maritima* D54C/C81S CheY. b.) RP-HPLC after 3 hr of thrombin digestion. Two peaks appear in the preparation of O pool 2 *T. maritima* CheY even before thrombin digestion. Cleavage of the his tag produces a protein that is slightly more hydrophobic, eluting 2 and 3 min later.

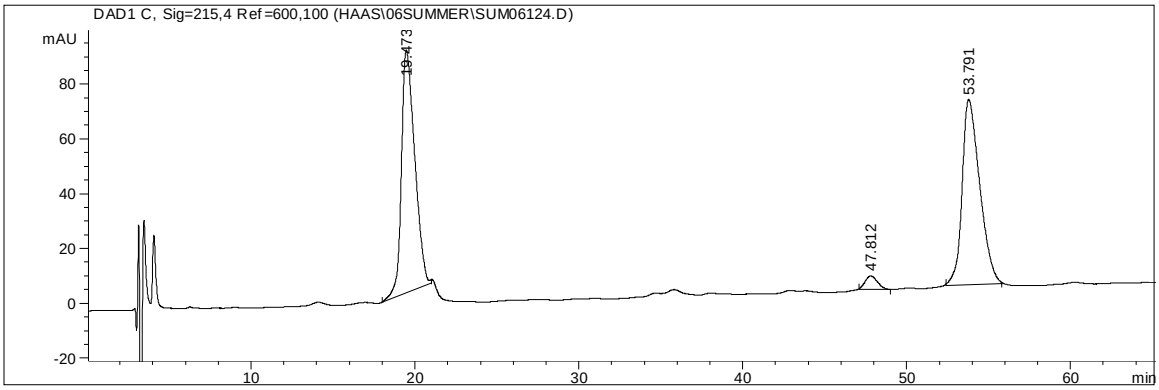


a.

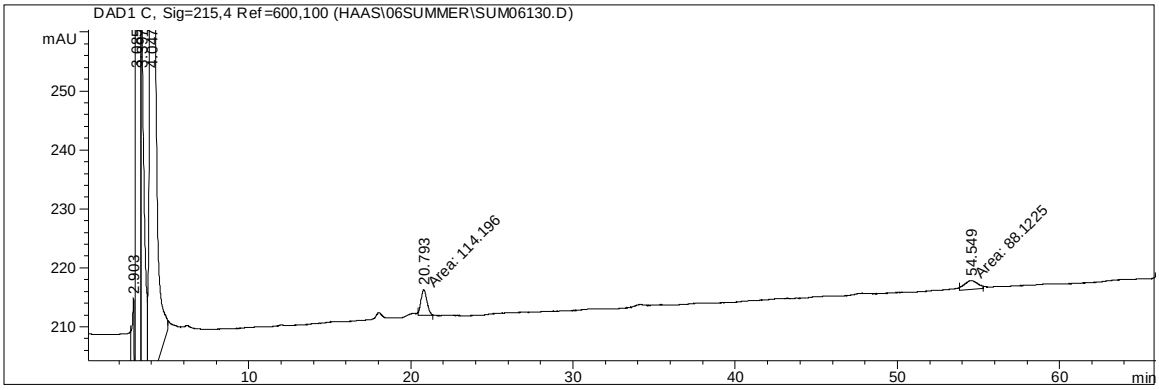


b.

Figure 48. RP-HPLC after 5 hr (a) and 23 hr (b) of thrombin digestion of *T. maritima* D54C/C81S CheY. The two peaks before cleavage of the his-tag elute at 7 and 32 min. After cleavage, the resulting protein elutes at 9 and 36 min. Cleavage is almost complete after 5 hr in agreement with SDS-PAGE. An additional species is found in the 5-hr time-point eluting at 38 min.



a.

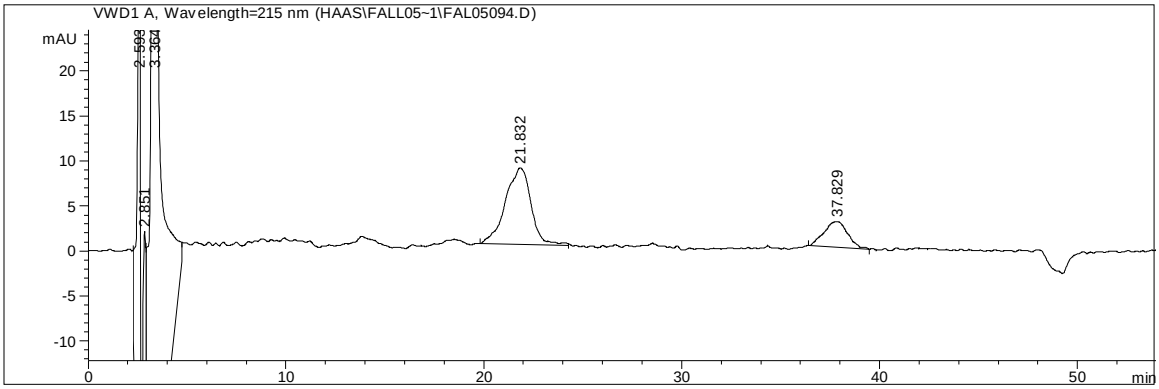


b.

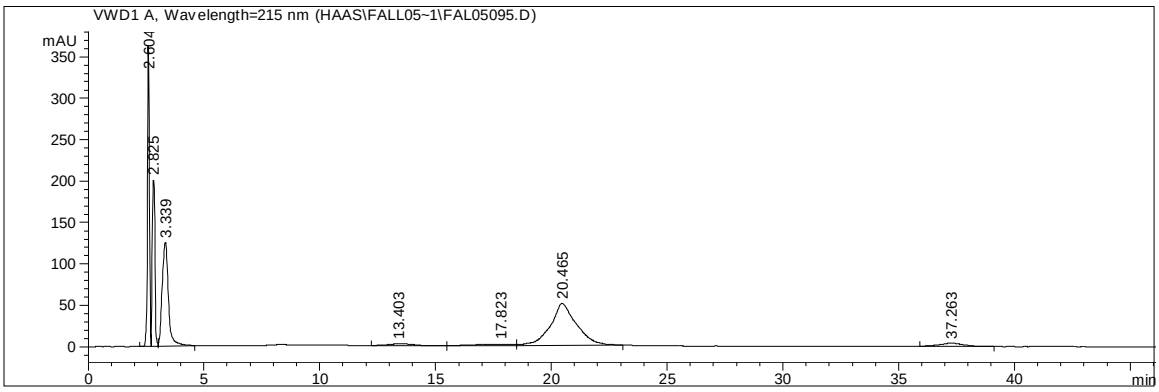
Figure 49. a.) RP-HPLC of the low salt fraction from the DEAE anion-exchange column. b.) RP-HPLC of the high salt fraction from the DEAE anion-exchange column. The two peaks present in UNCW preparations of *T. maritima* D54C/C81S CheY are found in both washes of a pH 8.6 anion exchange column.

	ε_{205}	$A_{280}^{0.1\%}$	ε_{280}
BSA	29.69 (29.7)[52]	0.663 (0.67) [52]	44270 (42900) [52]
Egg Albumin	30.06 (--)	0.763 (0.72)[53]	22830 (31700) [53]
TMY2'	27.82 (--)	0.19 (--)	2500 (2560)

Table 6. Calculated Molar Absorptivities. Data appears as Calculated Value (Literature Value).

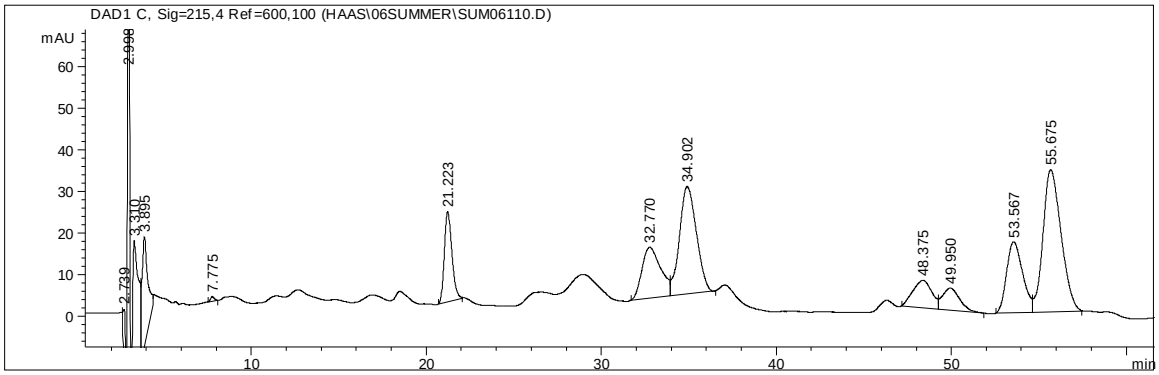


a.

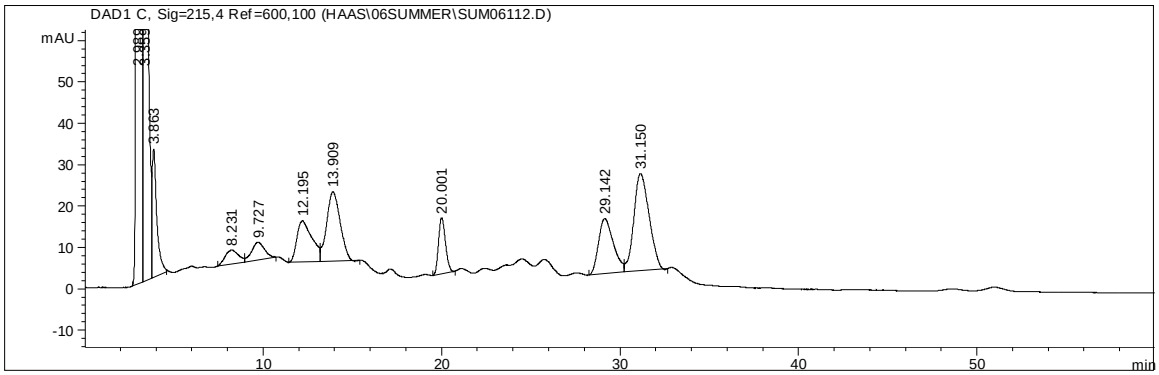


b.

Figure 50. a.) RP-HPLC of putative purified phosphono-CheY from *T. maritima* (15’’’). b.) RP-HPLC of 15’’’ spiked with CheY biotinylated with PEO-iodoacetyl biotin, the same biotin reagent used for purification. A spike experiment was performed in order to determine if the 22 min peak in panel a was purified phosphono-CheY or if it contained two species. No conclusions could be made as to the identity of the 22 min peak, possibly because the spike was too large an amount (or not enough) compared to phosphono-CheY. Close to a 1:1 ratio of the two species would be necessary to resolve them by RP-HPLC.

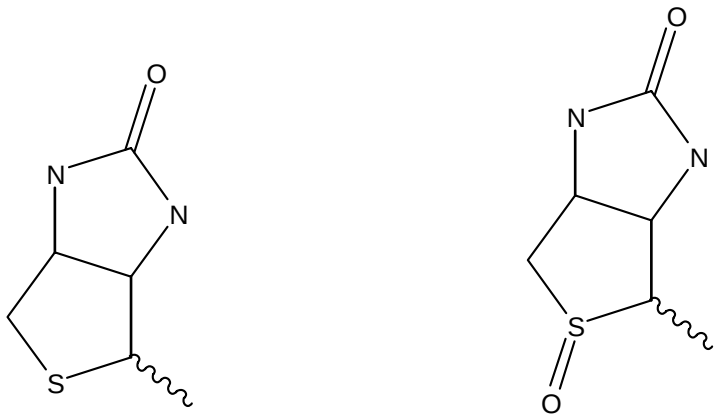


a.



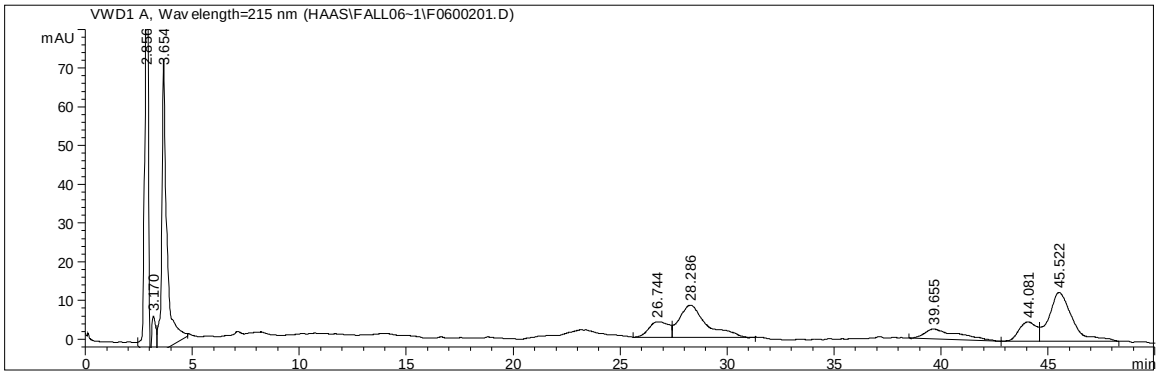
b.

Figure 51. a.) RP-HPLC of a *T. maritima* D54C/C81S CheY phosphonomethylation reaction mixture. b.) RP-HPLC of the same sample after biotinylation with PEO-iodoacetyl biotin. Unmodified CheY elutes as a doublet at 55 min and phosphono-CheY elutes as a doublet at 33 min. The appearance of the 12 and 13 min peaks in the biotinylation mixture (panel b) indicates heterogeneity resulting from the biotinylation reaction. Close examination of the phosphonomethylation mixture in panel a shows a similar pattern of peaks in the 12-14 min region, although they are much less pronounced relative to the other peaks in the chromatogram.

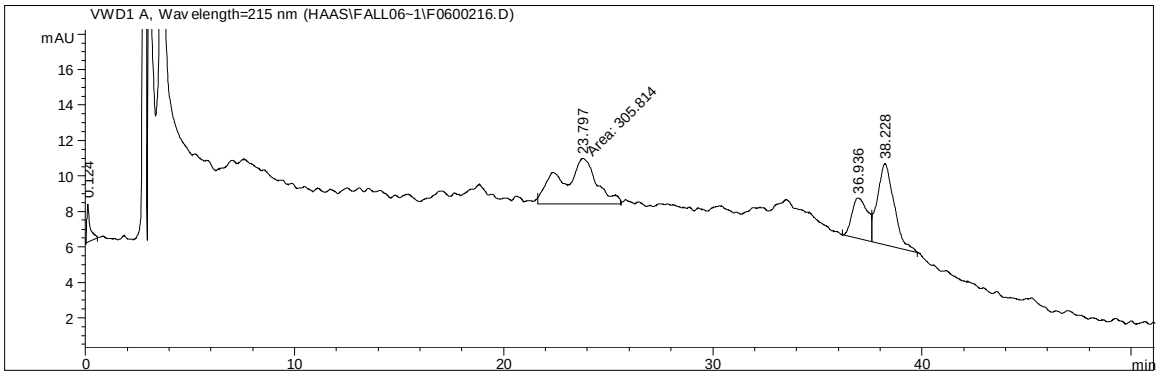


a. b.

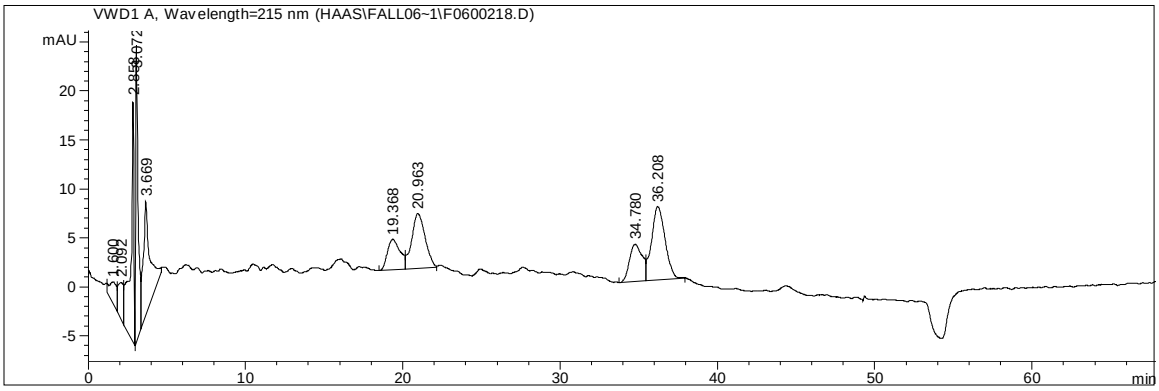
Figure 52. Biotin in its reduced (a) and oxidized (b) forms. The oxidized form of biotin may change the retention time of biotinylated *T. maritima* D54C/C81S CheY. It may also possess less affinity for avidin, making removal of biotinylated CheY more difficult.



a.



b.



c.

Figure 53. a.) RP-HPLC of a phosphonomethylation reaction mixture of *T. maritima* D54C/C81S CheY. Phosphono-CheY elutes as a doublet at 27 and 28 min, unmodified CheY elutes near 45 min. Putatively oxidized CheY is present at 39 min. b.) RP-HPLC of the same phosphonomethylation mixture after reaction with biotin-HPDP. CheY labeled with biotin-HPDP elutes as a doublet at 37 and 38 min. Phosphono-CheY now elutes at 22 and 23 min, presumably. c.) RP-HPLC of the same sample after the avidin column. Retention times are shifted earlier than in panel b, but the difference between the two remains the same.

DISCUSSION

Phosphonomethylation of *E. coli* CheY and Purification of Phosphono-CheY

Phosphonomethylation of *E. coli* D57C CheY by the standard protocol described generally proceeded to greater than 50% completion by DTNB and RP-HPLC. Protein recovery after the reaction was ~75% efficient and losses were attributed to precipitation and sample handling during buffer exchange and concentration. Protein precipitation within the phosphonomethylation reaction mixture was most likely a result of the alkaline conditions necessary to maintain cysteine nucleophilicity while minimizing the reactivity of lysine residues, whose amino group has a pK_a near 10.3. Protein precipitation may also be caused by the organic base triethylamine. When the inorganic base KOH was used as the proton scavenger precipitation as well as protein loss was minimal. The precipitation observed during reaction may not have been exclusively insoluble protein but may also have been formed from the metal ion reacting with hydroxymethylphosphonate, the product formed upon reaction of the triflate leaving group with water. Once the precipitate was removed by centrifugation, purification of phosphono-CheY from the reaction mixture by biotinylation and avidin was sufficient and generally yielded 30% of the total amount of protein reacted with PMT.

Purification of phosphono-CheY from unmodified *E. coli* D57C CheY using biotinylation and avidin was robust but required a significant amount of time. Biotinylation was allowed to proceed overnight (~18 hr) even though RP-HPLC and DTNB indicate the reaction with PEO-iodoacetyl biotin was 99% complete by 5 hours (Table 2). These results are consistent with those obtained earlier (Tyler Davis, Honors Thesis), suggesting that the reaction does not require overnight incubation. This provides a possible means to shorten the total time of purification, but only if one is careful to ensure completeness of biotinylation before taking the

sample over the avidin column. Biotin-HPDP, which was used exclusively with *Thermotoga maritima* protein, allows for the biotinylation reaction to be monitored spectrophotometrically, which may be the most time-effective method to ensure completeness of biotinylation.

The greatest time-dependent bottle-neck in the purification of phosphono-CheY is the avidin column. Due to the low binding capacity of monomeric avidin (~1 mg biotinylated protein per 5 mL settled gel; Pierce literature), two 5 mL columns are required to capture and retain biotinylated CheY for sufficient purification of phosphono-CheY. The use of two avidin columns, while effective in overall purification, increased the time required to obtain the sample in a form adequate for storage (~300-500 µL); often the time required to concentrate a volume of 30 mL to less than 1 mL would span two or more days. As a result, a greater amount of protein was lost to dilution and in quantitative transfer from two centriprep concentrators rather than one. Centriprep concentrators are efficient, but their membranes tend to retain protein making it difficult to wash the membrane without severely diluting the concentrate.

The biotinylation/avidin protocol produced samples of purified *E. coli* phosphono-CheY that were 90% pure or better, but it was time-consuming and the resulting protein loss dropped the final yield to ~30% of the starting material; therefore, an alternate method for purification was explored. It was initially thought that when applied to a cibacron blue column, phosphono-CheY would elute under milder conditions than unmodified CheY, and this might provide an alternative method for separating the two species from a reaction mixture. It appeared that the blue column was unable to completely differentiate between phosphono-CheY and unmodified CheY, since both species were present in the low salt and high salt fractions. However, RP-HPLC analysis indicated that unmodified CheY preferentially eluted in the high salt wash suggesting that partial purification of phosphono-CheY from unmodified CheY could be

achieved. In addition, protein quantitation of the fractions eluted from the cibacron blue column revealed that the majority of protein had to be eluted with 2 M NaCl, suggesting that both protein species had some affinity for the column. Further experiments may identify a set of buffer and salt conditions in which phosphono-CheY elutes and unmodified CheY remains bound to the column. Two columns used in tandem may also be useful if one column is only able to achieve partial purification. The column could also be used to separate out the 11 minute peak if this were an unwanted contaminant such as di-phosphonomethylated CheY.

Production and Purification of *T. maritima* D54C/C81S CheY

Successful preparations of *T. maritima* D54C/C81S CheY minimize the time from cell lysis to final purification and include 50 mM 2-ME in the protein solution after the size-exclusion column. Including the large 2 L growth of cells and assuming 2 days for concentrating, the protein preparation can be completed in 5 days. The major time constraint is dialysis of the Ni-NTA eluate into thrombin cleavage buffer, which takes a minimum of three days after which, as a result of the consequent protein dilution, the sample must be concentrated to a volume suitable for loading onto the SEC column (<3% of column volume). It was experimentally determined that dialysis into thrombin cleavage buffer was not necessary and that the addition of 5 mM $MgCl_2$ and 2.5 mM $CaCl_2$ directly to the Ni-NTA eluate is sufficient for complete thrombin digestion. This reduced total preparation time by 3 days (approximately in half) and eliminated protein precipitation within the dialysis.

In preparations that were dialyzed prior to thrombin digestion, precipitation was always found in the dialysis bag; moreover, it remained during all subsequent steps (concentration, digestion, and loading of the gel filtration column). As a result, early preparations of

Thermotoga maritima D54C/C81S CheY had to be syringe filtered onto the gel filtration column to prevent it from clogging. The lack of visible precipitation in preparations digested directly in the Ni-NTA eluate could be a result of imidazole, which is present at 200 mM in the elute buffer and may help stabilize the protein and prevent aggregation with his-tag peptide. More likely, though, is the remaining 5% of protein impurities present after purification by IMAC (see Results) lead to non-specific protein aggregation and contribute to CheY precipitating out of solution. Furthermore, this aggregation may be exacerbated by the increase in time required for dialysis, consistent with lack of visible precipitation if the dialysis and concentration steps were removed. Elimination of dialysis appears absolutely necessary for successful and efficient production of pure *T. maritima* D54C/C81S CheY, and this has not posed a problem since adequate cleavage of the his-tag is achieved with 1:2000 w:w thrombin:protein in the presence of 200 mM imidazole, 2.5 mM $CaCl_2$, and 5 mM $MgCl_2$.

The other key to successful preparations of *T. maritima* D54C/C81S CheY is the addition of 50 mM 2-ME to the fractions that contain CheY from the gel filtration column. In contrast, the addition of 10 mM DTT as suggested by the Cornell protocol may prevent oxidation, but it failed to keep the protein in solution during concentration. 50 mM 2-ME not only prevented oxidation and precipitation, but it also led to a more homogenous protein sample (lacking CheY*) as analyzed by RP-HPLC.

Phosphonomethylation of *T. maritima* D54C/C81S CheY

The majority of *T. maritima* phosphonomethylations were carried out in 250 mM AMPSO, pH 9.01 with 125-200 mM Sr^{2+}, Ca^{2+} or Ba^{2+} (chloride salts), 120 mM PMT and three equivalents of triethylamine. 250 mM TAPS, pH 8.50 was also frequently used and did not give

significantly different results. Sr^{2+} gave slightly greater yields, however, Ba^{2+} gave greater consistency. Nd^{3+} and Lu^{3+} at 10 mM generated alkylations as good as the divalent metals, although they were never as clean by RP-HPLC. The final concentration of these metals was lower than 10 mM due to the low solubility product constants of their hydroxide salts (3.2×10^{-22} and 1.9×10^{-24}, respectively). Since *T. maritima* proteins do not like high ionic strength conditions (Will Deutschman, Private Communication), CheY was kept to a low concentration (5-8 mg/ml) in alkylations with 200 mM divalent metal and was increased to 10-14 mg/mL in alkylations with 125 mm divalent metal or with the trivalent metals. Increasing the pH of the buffers did not result in increased reaction yield: AMPSO at pH 9.25 and CAPSO at pH 9.6 and 10.1 were tried, and neither produced better results than TAPS at pH 8.5 or AMPSO at pH 9.0. However, alkylations done in 1.3 M AMPSO at pH 10.1 with 3 equivalents of KOH (instead of triethylamine) produced excellent yields between 60-70%, although these reactions were not clean by RP-HPLC. Phosphonomethylations carried out at 37 °C produced clean reactions although the yield appeared to suffer. Increasing the temperature to 50 °C did not increase the yield or homogeneity of the reaction. Phosphonomethylation of *T. maritima* D54C/C81S CheY were consistently in the 35-50% range, almost irrespective of the conditions. Every variable was changed at least once, and virtually every condition appeared to produce a good reaction at least one time.

The extent of phosphonomethylation was difficult to determine with any precision due to the disagreement in the DTNB assay and RP-HPLC peak area ratios: even in the cleanest-looking reactions by reversed-phase analysis, rarely did the assays agree within 10%. Discrepancy between the two analytical techniques was minimized by incubating the cuvettes for the DTNB reaction at 25 °C, which increased the rate of the reaction and removed any variability

from changes in ambient room temperature. However, as discussed before, reversed-phase analysis of phosphonomethylation reactions often resulted in more than two peaks and made identification of the phosphono-CheY peak difficult (for examples see Figures 17 and 19). The cause of this heterogeneity remains speculative, but several ideas have been posed. Various peaks may arise in RP-HPLC of proteins as a result of slight differences in protein conformation at the hydrophobic foot (or contact region between protein and the stationary phase). Proline residues are often found at this contact surface, and inter-conversion between the *cis* and *trans* conformations of these residues may be slow on the time scale of the chromatography, leading to the appearance of peak doublet patterns (Andy Alpert, private communication). A fairly simple experiment can disprove this hypothesis: collect one of the peaks of the doublet, let it stand for an hour or so to allow proline *cis-trans* inter-conversion, and re-inject the peak onto RP-HPLC. If the hypothesis were true, the doublet pattern should reappear.

Methionine oxidation can also give rise to peak doubling and may contribute to the heterogeneity seen in RP-HPLC analysis of phosphonomethylation reaction mixtures. When methionine is oxidized to methionine sulfoxide, one oxygen atom is added to the thioether sulfur atom, and diastereoisomeric methionine sulfoxides are formed in nearly equal amounts (Lowell Ericsson, ABRF discussion). These chromatograph as a doublet that is not always symmetrical. Several attempts to disprove this hypothesis were unsuccessful (i.e. did not reduce heterogeneity or peak doubling). 2.5 M urea with 10 mM DTT had no effect on the peak doubling, nor did 10 mM BMS or 100 mM DTT. A systematic investigation of the reduction of methionine sulfoxide to methionine in peptides and proteins (*54*) describes a procedure involving 720 mM DTT in 5% acetic acid at 37 °C. This also failed to reduce the peak doubling, although these results are less clear since 75% of the protein was lost to precipitation in the low pH conditions. The next step

to disprove methionine oxidation is to use the novel reducing agent N-methylmercaptoacetamide. N-methylmercaptoacetamide has to be synthesized according to (*54*), but it has been shown to selectively reduce methionine sulfoxide back to methionine and can be used to test methionine oxidation as the cause for the heterogeneity and peak doubling seen in RP-HPLC analysis of phosphonomethylation reactions.

The heterogeneity observed by RP-HPLC of phosphonomethylation reaction mixtures may not be a result of proline isomerization or methionine oxidation, but may be the result of multiple phosphononomethyl modifications of CheY (for example, at lysine or histidine residues). Several mass spectrometry experiments suggest this is a plausible explanation. The mass spectrum of phosphono-CheY (Figure 39) and the mass spectrum of the RP-HPLC peaks prior to phosphono-CheY (Figure 38) indicate species of masses consistent with one, two, three, and four phosphonomethyl modifications to *T. maritima* D54C/C81S CheY. The presence of these species was confirmed by MS analysis of the trypsin digests as well as two additional intact protein MS experiments. The first experiment was a minute-by-minute peak collection from RP-HPLC of a phosphonomethylation reaction mixture followed by direct injection of each fraction into an ESI-TOF mass spectrometer. The second experiment was a gradient LC-MS of the same reaction mixture (i.e. intact protein mixture analyzed on the same shallow RP-HPLC gradient with a time-of-flight MS detector rather than a variable wavelength detector). These experiments demonstrated not only that multiple phosphonomethylated CheY proteins were present, but also that their reversed-phase retention times were negatively correlated with the number of phosphonomethylations (+94 Da). Phosphonomethylation reduces hydrophobicity, and the protein has a greater ability to partition into the mobile phase when it has more phosphonomethyl modifications.

There exist a number of pieces of evidence that suggest *T. maritima* D54C/C81S CheY is modified not only on the engineered Cys54 but also on two Lysine residues. In order to test this idea, attempts were made to reduce the nucleophilicity of the Lysine residues by decreasing the pH at which the phosphonomethylation reaction was performed. Phosphonomethylations at pH 7.5 and 8.0 in 250 mM HEPES and TAPS did not produce any modified CheY species by RP-HPLC. Reactions carried out with reduced amounts of PMT (60 mM) were also unable to produce a single additional species by RP-HPLC. The presence of multiple peaks in reversed-phase analysis of the reaction mixture suggests but does not demand that more than one residue has been modified. Only tentative conclusions as to what and how many new species were formed in each reaction can be made by RP-HPLC analysis since definite peak assignments have not been made from mass spectrometry analysis. This has not yet been possible because mass spectrometry analysis indicates there is more than one protein species in each individual reversed-phase peak. For the gradient LC-MS experiment of the phosphonomethylation reaction mixture, the tubing between the column and the electrospray source was much too long and led to peak broadening. As a result, all that was seen in the total ion count chromatogram was a modest increase in the baseline, and no real discernable "peaks" were resolved. The overlap in protein species, however, is hard to explain in the MS analysis of the minute-by-minute reversed-phase fractions.

In order to make proper and conclusive RP-HPLC peak assignments, mass spectrometry analysis must be done in such a manner that no species overlapping occurs. Each peak needs to be collected individually, lyophilized to remove TFA, and analyzed by direct syringe infusion electrospray MS (rather than direct *injection* using the Agilent 1100 autosampler). The protein can be brought up to between 1 and 10 µM in 50:50 ACN/Water with 0.1% formic acid. The

direct infusion MS analysis should be performed at the slowest flow rate possible (50 µL/min maximum) and the volume of the sample should be such that spectrums can be obtained continuously for 5 minutes to give adequate signal-to-noise. One should determine the ideal protein concentration for this MS experiment by preparing a range of myoglobin standards (1, 2, 5, and 10 µM) in the same solution as the protein of interest will be dissolved in. Myoglobin has a molecular weight similar to CheY and should give reasonably similar responses in electrospray mass spectrometry (Will Thompson, Private Communication).

Biotinylation of *T. maritima* D54C/C81S CheY

Kinetic Treatment

A detailed kinetic analysis was performed on the pseudo first-order reaction of biotin-HPDP with unmodified *T. maritima* CheY and half-lives ranging from 9.5 minutes to 17.6 minutes were calculated. Data were obtained spectrophotometrically by following absorbance at 343 nm, a measurement directly proportional to the amount of pyridine-2-thione released in the reaction vessel. Absorbance values taken over a 3 hour period were analyzed three ways in order to determine an upper-bound estimate of the 10^{th} half-life of the reaction (99.9% completion). The first was linear least-squares fit to the first-order rate equation; the second was exponential regression analysis using the logarithmic form of the first-order rate equation; the third was the Guggenheim method (*55*). These three methods were satisfactory since the reaction was set up to follow pseudo first-order kinetics, but ideally one would first use the method of initial rates to determine the order of the reaction before proceeding with the analysis.

Each data analysis method had its advantages and disadvantages. All three methods were fairly simple, although the Guggenheim was a little more difficult. The linear form of the first-order rate equation required an estimate of the endpoint of the reaction, which was difficult to

determine due to side-reactions (possibly hydrolysis) that released pyridine-2-thione at a slow but observable rate. For this technique, absorbance at the endpoint was estimated by the theoretical ΔA_{343} calculated from the concentration of free thiol determined by DTNB, since it was this free thiol that would react with the biotin reagent, releasing a molar equivalent of the chromophore. The other two techniques did not require a value for the endpoint of the reaction which was advantageous. The Guggenheim method had one disadvantage, though: it can give linear plots for reactions that are second-order so the reaction should be a known first-order reaction or the data can be misleading. Assuming that pseudo first-order kinetics holds in this experiment, the reaction with Biotin-HPDP has an upper-bound estimate for the 10^{th} half-life of 3 hours.

Biotinylation Prior to Purification

Biotinylation of unmodified *T. maritima* D54C/C81S CheY with all three biotin reagents produces essentially a single species by RP-HPLC. CheY labeled with biotin-HPDP produced a species that eluted between phosphono-CheY and unmodified CheY. CheY labeled with PEO-iodoacetyl or PEO-maleimide eluted ~1.5 and 2.5 min earlier, respectively, than phosphono-CheY by RP-HPLC. Chromatographic resolution calculations (*56*) indicate that both CheY biotinylated with PEO-iodoacetyl and CheY biotinylated with PEO-maleimide can theoretically be resolved from phosphono-CheY at half the peak height; in practice, however, this was not the case. Spike experiments, in which pure biotinylated protein was added to a double-prime sample, were aimed at identifying the biotinylated CheY peak but were not successful (Figure 50). Monitoring the disappearance in the unmodified CheY peak in RP-HPLC was the only way to determine the extent of biotinylation with PEO-iodoacetyl and PEO-maleimide biotin, since their retention times nearly overlap that of phosphono-CheY. Moreover, distinguishing

phosphono-CheY from biotinylated CheY was nearly impossible when phosphono-CheY eluted as a doublet in the reversed-phase analysis of phosphonomethylation reaction mixtures (see Figure 21). Biotin-HPDP was advantageous in that the presence of biotinylated CheY could easily be confirmed by RP-HPLC.

Biotinylation of unmodified CheY is a clean reaction producing a single species by RP-HPLC as discussed before. However, evidence of heterogeneity in a phosphono-CheY mixture after biotinylation with PEO-iodoacetyl was found (Figure 51). The peaks elute much earlier than supposed biotinylated-CheY and may be the result of oxidation of the biotin ring (Figure 52), although no evidence of this was found in the biotinylation of no-prime samples. The unknown peaks remained throughout purification, suggesting that either they are not biotinylated CheY, or that the biotin group, in its oxidized form, shows little affinity for avidin. It is hard to determine if the peaks appear as a result of biotinylation or if they were already present in the phosphonomethylation. The RP-HPLC analysis of the phosphonomethylation mixture (Figure 51a) indicates two small peaks in this vicinity; however, they are much less intense. This could be the unlikely result of the preferential precipitation of phosphono-CheY and unmodified CheY during reversed-phase analysis, but it may also be different populations of biotinylated protein. Mass spectrometry indicates that more than one protein species may be produced during phosphonomethylation, and if this is the case, biotinylation of these species will also likely produce more than one unique protein. Due to the specificity of the biotinylation reagents, this is possible only if the species are CheY proteins that have been modified at a residue other than Cys54. Evidence from HIC experiments also suggests that more than one protein species is reactive towards biotin. Tandem LC-MS/MS is the definitive method for identifying these species and their site of modification.

Anomolous results were also observed in the biotinylation of unmodified CheY with biotin-HPDP (Figure 53). Biotinylation with biotin-HPDP should produce only one species, however two peaks were observed upon biotinylation of a phosphonomethylation reaction mixture (Figure 53b). This result could be explained by conformational heterogeneity of a single species of biotinylated CheY or by a chemical change to CheY that rendered two different species of CheY available for the biotinylation reaction. The presence of doublets at the position of CheY and phosphono-CheY before biotinylation (Figure 53a) suggests the latter but does not rule out conformational heterogeneity. In addition to the possibility of two biotinylated species, odd results were found in the RP-HPLC retention time of phosphono-CheY, which is almost 10 minutes earlier than normal while the retention time of biotin-HPDP does not change (Figure 53b). The peaks may represent a new species altogether (for instance, CheY biotinylated at Cys-54 and phosphonomethylated at Lys 67), but putatively, the peak represents phosphono-CheY only it elutes earlier. This may be the result of some interaction between phosphono-CheY with biotin-HPDP-labeled CheY that disrupts the interaction between phosphono-CheY and the stationary phase. Gradient LC-MS would be able to definitively determine that the earlier-eluting species is still phosphono-CheY.

Purification of *T. maritima* Phosphono-CheY

Biotin/Avidin Protocol

Purification of phosphono-CheY has been achieved with limited success. The most common problem has been the incomplete removal of biotinylated CheY by the avidin column. An adequate explanation for remains elusive since the biotin-avidin complex is one of the strongest known; biotin has a dissociation constant for monomeric avidin and tetrameric avidin

as supplied by Pierce of 10^{-8} M^{-1} and 10^{-15} M^{-1}, respectively. A number of experiments failed to resolve this problem. Overnight incubation of the protein on the column did not remove biotinylated CheY, nor did changing the buffer and pH at which the column was run. The use of high affinity tetrameric avidin also was unsuccessful in removing biotinylated CheY. These experiments all used biotin-HPDP as the CheY-labeling reagent, so the presence of biotinylated CheY was easily observed by RP-HPLC. One would not be able to exclude the presence of biotinylated CheY by RP-HPLC if PEO-iodoacetyl or PEO-maleimide biotin were used. This sheds doubt on the purity of any phosphono-CheY sample in which these water soluble reagents (PEO-biotin) were used, as the reversed-phase peak putatively identified as phosphono-CheY may not be a single species.

The second problem encountered during the purification of *T. maritima* phosphono-CheY was the apparent re-emergence of the unmodified CheY peak after avidin chromatography. This anomalous result was regularly found if unmodified CheY was labeled with PEO-iodoacetyl biotin, suggesting either that the thioether bond is labile, or there is a secondary site of modification in unmodified CheY. When PEO-maleimide biotin was reacted with unmodified CheY, the result appeared to be purified phosphono-CheY. Mass spectrometry of these samples would identify any species that co-elute in the putative reversed-phase phosphono-CheY peak.

Immobilized Glutathione

An alternate method for the purification of phosphono-CheY uses a reduced glutathione column as previously published (*41*). Glutathione is a naturally occurring, cysteine-containing tri-peptide used for the reduction of peroxides within the body. Reduced glutathione binds and

reacts with disulfide bonds, reducing and displacing one of the species and forming a new disulfide with the other. For our purposes, a mixed disulfide of CheY (TNB-labeled CheY) must react with the immobilized glutathione, bind to the column, and release the TNB anion. Only 35% of the TNB-labeled protein reacted with the glutathione and was eluted with reducing agent. The rest of the protein did not bind to glutathione and was washed off easily. The most reasonable explanation for this result is that the mixed disulfide between glutathione and TNB formed instead of the disulfide between glutathione and unmodified CheY. Analysis of the wash by RP-HPLC supports this conclusion. The position of the peak by RP-HPLC is consistent with the position of unmodified CheY, not TNB-labeled CheY. Derivatization of the glutathione with DTNB prior to application of the phosphonomethylation reaction mixture may produce better results.

Hydrophobic Interaction Chromatography

HIC-HPLC provides a non-denaturing method for the analysis and possible purification of *T. maritima* phosphono-CheY from a reaction mixture. RP-HPLC analysis of the fractions collected from HIC-HPLC of a phosphonomethylation mixture indicated that phosphono-CheY can be found between 13.5-17 minutes. This is consistent with elution of phosphono-CheY as identified through MALD/I-MS analysis of previous HIC collections (Figures 43 and 44).

The multiplicity of peaks found in the first 13.5 minutes of HIC-HPLC of a phosphonomethylation reaction mixture are likely related to the heterogeneity found in the RP-HPLC analysis of a phosphonomethylation mixture. MS analysis indicates oxidation both within the phosphono-CheY peak and in the peaks eluting immediately prior to phosphono-CheY. The peaks present in HIC-HPLC of a phosphonomethylation reaction mixture are likely reflecting

these various oxidation states, particularly since HIC is much more sensitive to positional effects than RP-HPLC (Andy Alpert, personal communication). HIC retains the tertiary structure of a globular protein; therefore, the stationary phase is only able to interact with the hydrophobic residues at the surface of the protein. Reversed-phase on the other hand is able to disrupt the tertiary structure of a protein through solvation of the core hydrophobic residues, enabling all of these residues to interact with stationary phase. HIC is therefore much more sensitive to polarity changes of single amino acids than is reversed-phase.

The peaks present in HIC-HPLC of a phosphonomethylation reaction mixture not only support the oxidation hypothesis, but also the hypothesis of multiple phosphonomethylated CheY species. The first evidence that suggested HIC-HPLC was separating differentially phosphonomethylated proteins was found from the attempts to disprove methionine oxidation (see Discussion above of Phosphonomethylation of *T. maritima* CheY). The conditions used to reduce methionine sulfoxide were unable to "clean-up" not only the RP-HPLC but also the HIC-HPLC of a phosphonomethylation reaction mixture. The second piece of evidence suggesting that HIC-HPLC was separating differentially phosphonomethylated proteins was found in the analysis of a PEO-iodoacetyl biotinylation. It was known that biotinylation of *T. maritima* D54C/C81S CheY with PEO-iodoacetyl biotin produced a species that eluted ~4 minutes after unmodified CheY (see Figures 28a and b). When a phosphonomethylation reaction mixture was subject to biotinylation with PEO-iodoacetyl biotin, not only is the CheY peak shifted back 4 minutes but three other peaks appear to elute later as well. This suggests that there is more than one CheY species with a reactive cysteine, assuming that PEO-iodoacetyl biotin remains specific to thiols at pH 8.5.

Characterization of CheY*

In the preparation of the *Thermotoga maritima* double mutant D54C/C81S CheY, the protein is bound by a hexa-histidine tag to a chelating Nickel column and eluted with an increased concentration (200 mM) of imidizole. The his-tag is removed by thrombin digestion, and the resultant mixture separated by size-exclusion chromatography. All of the protein preparations done in our lab (except the recent preparations A-D) produced pure protein by SDS-PAGE and Coomassie Blue staining but contained two peaks when analyzed by RP-HPLC. A number of experiments aimed at identifying this impurity (CheY*) were performed.

To test whether CheY* is his-tagged CheY, a time course of a thrombin digestion was followed to completion both by SDS-PAGE and RP-HPLC. The presence of two peaks before thrombin digestion with nearly the same retention times as the two peaks seen in RP-HPLC chromatograms of UNCW CheY preparations M through T immediately suggests that CheY* is not his-tagged CheY. Moreover, the CheY* peak remained throughout the analysis even though SDS-PAGE indicated thrombin digestion was complete by 23 hours. A puzzling result in this experiment is the decrease in the early-eluting peak at the 5-hr timepoint, suggesting CheY* may in fact be his-tagged CheY. The eventual increase of the peak in the 23-hr sample is evidence to the contrary, but what can explain the difference in peak area ratios of the early and late-eluting species between time 0 and 23 hours? Evidence of a time-dependent fluctuation in apparent peak percentages has been observed at other times, but no concrete explanation can be found. Paying close attention to the buffer conditions and reduction state of the protein will help describe this time-dependent fluctuation.

Disregarding the change in peak area percent in the 5-hr and 23-hr samples and the appearance of a third peak eluting after CheY, thrombin digestion appears to affect both peaks in

relatively the same manner. The newly cleaved protein has slightly greater hydrophobicity than its parent, which is consistent with removal of the 6-His tag, a short and mostly hydrophilic 16 amino acid peptide. The tag has 7 hydrophilic amino acids (6 His, 1 Arg), 6 neutral amino acids (2 Gly, 3 Ser, 1 Pro), and only 3 hydrophobic amino acids (1 Met, 1 Val, 1 Leu) and would theoretically create a protein more hydrophobic upon its removal. A second experiment was also performed, in which a heterogeneous sample (50:50 CheY*:CheY) was passed back over the Nickel column. No protein bound to the column, and the peak did not disappear by HPLC, which is further evidence indicating CheY* is not his-tagged CheY.

The possibility that CheY* was different in primary sequence was investigated. If it contained additional amino acids it may be due to ribosomal read-through during translation of the mRNA. If it contained fewer amino acids it could be created by undesired proteolysis. Digestion produces 14 peptides of varying hydrophobicity and ranging in mass from 334 Da to 2204 Da. Broad gradient RP-HPLC analysis of the digestions of O Pool 1 and of Cornell Thermotoga CheY with endoproteinase LysC showed impressive consistency with respect to relative peak size and overall shape of their chromatograms. The only difference observed was an increase in peak height around 33.3 min for O Pool 1. One would expect to see a new peak in one of the samples if the proteins were different, since each peptide is produced in the same molar quantity. There must be an extra peptide in UNCW CheY that co-elutes with the 33.5 min peptide, indicating the UNCW sample may not be identical to the Cornell sample at the amino acid level. Is this a result of differential thrombin cleavage or by ribosomal read-through at the level of translation? Could there be amino acid cleavage at the C-terminus? LysC digestion of his-tagged protein could at least rule out differential thrombin cleavage, but amino acid analysis would be necessary if one thought the sequences were different.

To test whether the peak was oxidized CheY, a sample that was 50% CheY* by HPLC was treated with 4 different reducing agents. Neither 22 mM 2-mercaptoethanol nor 5 mM TCEP, DTT, and BMS were able to decrease the area of this peak. TCEP, however, was dissolved in phosphate buffer which reduces its stability and may have limited its capacity for reduction (Technical literature). While these experiments were not able to get rid of the CheY* peak, they did identify oxidized CheY as eluting almost exactly 5 minutes before the CheY peak. Interestingly though, prep R was done entirely in 10 mM 2-ME and the size of the CheY* decreased to ~ 1/2 of its usual size (~20% by area). Another interesting observation is the CheY* peak does not oxidize. It may be present in too small a quantity to reliably see by absorbance at 215 nm, or it may elute in the void volume prior to the gradient starting conditions.

RP-HPLC Conditions Lead to Peak Heterogeneity in *T. maritima* CheY

The unknown species (CheY*) present in RP-HPLC chromatograms is most likely a non-specific aggregation of CheY molecules. Collection of CheY* and re-injection on RP-HPLC gives both CheY* and a small amount of CheY. The appearance by RP-HPLC did not change after letting the sample sit out for 24 hours. When run by HIC-HPLC, only a single peak was found even though reversed-phase indicated two species were present. Since HIC-HPLC is a native technique and RP-HPLC is presumed to be denaturing, the presence of two peaks by RP-HPLC indicates a conformational issue unique to the reversed-phase conditions.

Quantitation of *T. maritima* D54C/C81S CheY

Calculation of the molar absorptivity of *T. maritima* CheY by the methods of Gill and von Hippel (*46*) gives 2560 M^{-1} cm^{-1}. *T. maritima* CheY, therefore, does not absorb much light at 280 nm, the standard wavelength used for protein quantitation. This was of some considerable

concern, as our protein concentrations are usually kept below 1.0 mM, and greater error is associated with lower absorbance values. Moreover, our collaborators published results in which they use both Bradford assay and reductant compatible/denaturant compatible assay (BioRad) to estimate a molar absorptivity of 6500 M^{-1} cm^{-1} for CheY (*38*). Before *T. maritima* CheY experiments were undertaken, it was necessary to properly quantify the amount of protein present in solution, either by pursuing the molar absorptivity at 280 nm (or some other wavelength), or by finding a method other than UV absorption.

Bradford and BCA assays are two colorimetric assays that are very popular. Bradford monitors the shift in absorbance to 595 nm upon protein binding to Coomassie Blue dye. BCA monitors the complex formed between protein and Cu^{2+} by measuring absorbance at 560. Both colorimetric assays use a series of standards with protein concentrations ranging from 0 to 50 µg/mL by which a standard curve can be made. One can then estimate the concentration of a protein solution by finding its absorbance on the standard curve. Two errors are associated with these methods: First, standard curves obtained are not always linear over a wide range of concentrations, especially for the Bradford assay. Second, different proteins behave differently under these color-changing conditions, so there is no guarantee the protein of interest will behave similarly to the protein used to create the standard curve. For example, Bradford assay of bovine serum albumin (BSA) and cytochrome *c* give apparent concentrations 2.1 and 2.5 times greater than their actual concentration (BioRad Technical Bulletin LIT33 Rev C). This problem can sometimes be overcome by creating two standard curves, each with a different protein. Egg albumin behaves well in the Bradford Assay and was regularly used for the standard curve. Despite the drawbacks, colorimetric assays have many advantages. They are fast and relatively inexpensive, compatible with a wide variety of reagents, and very sensitive.

Colorimetric assays are useful, but destructive. UV absorption at 280 nm is non-destructive and is the easiest, most accurate technique for measuring protein concentration. However, UV absorption of proteins also has its drawbacks. Matched cuvettes that do not absorb above 190 nm and proper referencing are essential to obtaining an accurate protein spectrum. Fortunately, the slope of a protein spectrum near 280 nm is often gradual and less error is introduced by the bandwidth of the light. This is not the case with measuring absorbance at 205 nm, which is on the steep slope of the absorbance maximum of the peptide bond at 192 nm. For this reason, a Cary double-beam spectrophotometer was used to estimate the molar absorptivity of *T. maritima* CheY. This instrument has a number of advantages over the lab's Pharmacia spectrophotometer. First of all it has two beams that constantly correct for changes in lamp intensity over time. Second, the double-beam eliminates differences between reference and sample cuvettes. Third, the Cary scans at a slower rate which increases the signal-to-noise ratio. Finally, the Cary allows a correction for stray light effects by allowing the user to set 100% transmittance and 0% transmittance before measurements are taken. Greater accuracy and precision were achieved in the quantitation of *T. maritima* CheY on the Cary spectrophotometer than could have been achieved on the single-beam Pharmacia.

The reported method for estimating a protein's molar absorptivity at 280 nm by simply summing the number of tryptophan, tyrosine, and cystine (the oxidized disulfide form of cysteine) within a protein of interest provides a good estimate of the molar absorptivity of *T. maritima* CheY. Each tyrosine is assigned a molar absorptivity of 1280 M^{-1} cm^{-1}, tryptophan 5690, and cystine 120 M^{-1} cm^{-1}. Thermotoga CheY possesses no tryptophan, no disulfide bonds, and only two tyrosine residues and has an estimated molar absorptivity of 2560 M^{-1} cm^{-1}. The results obtained from the A_{205}/A_{280} experiments performed on *T. maritima* CheY indicate that

2560 M^{-1} cm^{-1} is a good estimate. Moreover, the estimate of protein concentration using this molar absorptivity agrees with the protein concentration obtained from the Bradford Assay (1.4 mg/mL).

Final Thoughts and Future Experiments

Preparations of *T. maritima* D54C/C81S CheY will be homogenous by RP-HPLC as long as three steps are followed: thrombin digestion is done immediately after the Ni-NTA column, 10 mM 2-ME is added to the thrombin digestion, and 50 mM 2-ME is added after the size-exclusion column. Phosphonomethylation reactions of *T. maritima* D54C/C81S CheY will most likely continue to lack reproducibility, and therefore conclusive RP-HPLC peak assignments must be made. Ideally, each of the reversed-phase peaks represents a single species, and one of them is phosphono-CheY, but it may turn out that more than one species elutes within each peak. If this is the case, carefully controlled mass spectrometry experiments using direct infusion electrospray MS would have to be done to identify the *main species* in each RP-HPLC peak. Direct infusion procedures are outlined in the Discussion section Phosphonomethylation of *T. maritima* CheY. Ultimately, interpretation of RP-HPLC results will only be accurate if peak assignments can be made.

Experiments must also be targeted at identifying those peaks that represent CheY protein phosphonomethylated at sites other than Cys-54. There already exists strong mass spectrometry evidence supporting multiple phosphonomethylations. The hypothesis agrees with the observation that the ratio of free thiol is almost always higher in the DTNB assay than it is in the corresponding reversed-phase analysis; moreover, it is supported by the fact that the ratio of the putative phosphono-CheY peak to the unmodified CheY peak is much smaller in HIC-HPLC

than it is in the corresponding RP-HPLC analysis. Conclusive RP-HPLC peak assignments will also be useful in determining to what extent multiple phosphonomethylation is occurring, particularly since RP-HPLC is quantitative, while there exists conflicting opinions on how to correlate quantity of a species in solution with its ionization behavior in electrospray mass spectrometry.

Once RP-HPLC can be used to accurately determine the extent of multiple phosphonomethylation, reaction conditions might be manipulated with the goal of decreasing the production of these unwanted species. Lowering the pH of the buffer and dropping the concentration of phosphonomethyltriflate should theoretically reduce or eliminate multiple phosphonomethylation. Preliminary results at pH 7.5 and 60 mM PMT gave virtually no alkylation and phosphonomethylations at pH 8.0 and 60 mM PMT produced poor alkylations. The alkylation described by Tyler Davis (Honor's Thesis) showed promising results as far as yield is concerned and might be investigated further in order to decrease unwanted byproducts of the reaction. Cadmium and zinc have greater affinity for sulfur atoms (although they are smaller in ionic radius than Sr^{2+} and Ca^{2+}) and should give increased specificity during the phosphonomethylation reaction. But obviously one cannot think solely of phosphonomethylation without a means to purify it.

HIC-HPLC shows great promise as a purification tool, and collection of the 13.5-17 minute eluate is known to produce 75% pure phosphono-CheY. Anion-exchange at pH 7.9 was not useful in purification of phosphono-CheY; however, it could be useful if the silica-based column were replaced with a resin-based column that could be operated at a pH above the pI (8.68) of the protein. Cation-exchange might also be a useful tool for purification and has shown promise thus far. Purification of *T. maritima* phosphono-CheY by the biotin/avidin protocol was

successful on the microgram-level of protein, but it may not be useful for large-scale purification. This is in contrast to the purification of *E. coli* D57C CheY in which a single 5 mL monomeric avidin column removes between 0.5 and 1.0 mg of biotinylated CheY. The similarity in the molecular weights of the two CheY mutants suggests there would not be a significant difference in capacity of the avidin column for the biotinylated forms. Further experiments with biotinylation and avidin might use batch purification rather than purification over the avidin column. The advantage with batch purification is the ability to mix the biotinylated protein with monomeric avidin beads to allow maximal diffusion and binding.

In addition to batch conditions, solution conditions might be found empirically and through literature searching in which the tertiary structure of CheY is partially destabilized but avidin remains completely native. These conditions may promote mobilization of the biotin group and spacer arm (~28 Å) if its access to the avidin subunits is blocked or disrupted under normal solution conditions. Wild-type avidin has a melting temperature of 85.5 °C alone and 118.2 °C when in complex with biotin (*57*). Pierce technical support gives a melting temperature estimate of 132 °C in complex with biotin. Note that these estimates are for non-immobilized avidin. Immobilized monomeric avidin is also compatible with up to 8 M Urea and 3 M guanidinium-HCl (Pierce Technical Support). *T. maritima* CheY can be almost completely unfolded at room temperature in 3 M guanidinium-HCl (protein present at 20 µg/mL at pH 6.0 in 50 mM sodium phosphate)(*58*). 2.5 – 3.0 M guanidinium might be a good starting point to destabilize the protein, increase the mobility of the biotin group, and maximize the removal of biotinylated CheY by immobilized monomeric avidin chromatography.

REFERENCES

1. Bourret, R. B., and Stock, A. M. (2002) Molecular Information Processing: Lessons from Bacterial Chemotaxis, *J. Biol. Chem. 277*, 9625-9628.

2. Bruce Alberts, A. J., Julian Lewis, Martin Raff, Keith Roberts, Peter Walter (2002) *Molecular Biology of the Cell*, 4th ed., Garland Science.

3. Volz, K. (1993) Structural Conservation in the CheY Superfamily, *Biochemistry 32*, 11741-11753.

4. West, A. H., and Stock, A. M. (2001) Histidine kinases and response regulator proteins in two-component signaling systems, *Trends in Biochemical Sciences 26*, 369-376.

5. Matsumura, P., Rydel, J. J., Linzmeier, R., and Vacante, D. (1984) Overexpression and Sequence of the *Escherichia coli cheY* Gene and Biochemical Activities of the CheY Protein, *J. Bacteriol. 160*, 36-41.

6. Borkovich, K. A., Kaplan, N., Hess, J. F., and Simon, M. I. (1989) Transmembrane Signal Transduction in Bacterial Chemotaxis Involves Ligand-Dependent Activation of Phosphate Group Transfer, *Proc. Natl Acad. Sci. U.S.A. 86*, 1208-1212.

7. Parkinson, J. S. (1993) Signal Transduction Schemes of Bacteria, *Cell 73*, 857-871.

8. Zhu, X., Rebello, J., Matsumura, P., and Volz, K. (1997a) Crystal Structures of CheY Mutants Y106W and T87I/Y106W. CheY Activation Correlates with Movement of Residue 106, *J. Biol. Chem. 272*, 5000-5006.

9. Dahlquist, F. W. (2002) Amplification of Signaling Events in Bacteria, *Sci. STKE 2002*.

10. Mesibov, R., Ordal, G. W., and Adler, J. (1973) The Range of Attractant Concentrations for Bacterial Chemotaxis and the Threshold and Size of Response over This Range: Weber law and related phenomena, *J. Gen. Physiol. 62*, 203-223.

11. Silversmith, R. E., Guanga, G. P., Betts, L., Chu, C., Zhao, R., and Bourret, R. B. (2003) CheZ-Mediated Dephosphorylation of the *Escherichia coli* Chemotaxis Response Regulator CheY: Role for CheY Glutamate 89, *J. Bacteriol. 185*, 1495-1502.

12. Lukat, G. S., Stock, A. M., and Stock, J. B. (1990) Divalent Metal Ion Binding to the CheY Protein and Its Significance to Phosphotransfer in Bacterial Chemotaxis, *Biochemistry 29*, 5436-5442.

13. Halkides, C. J., McEvoy, M. M., Casper, E., Matsumura, P., Volz, K., and Dahlquist, F. W. (2000) The 1.9 Å Resolution Crystal Structure of Phosphono-CheY, an Analogue of the Active Form of the Response Regulator, CheY, *Biochemistry 39*, 5280-5286.

14. Stock, A. M., Mottonen, J. M., Stock, J. B., and Schutt, C. E. (1989) Three-dimensional structure of CheY, the response regulator of bacterial chemotaxis, *Nature 337*, 745-749.

15. Volz, K., and Matsumura, P. (1991) Crystal Structure of *Escherichia coli* CheY Refined at 1.7-Å Resolution, *J. Biol. Chem. 266*, 15511-15519.

16. Stock, A. M., Martinez-Hackert, E., Rasmussen, B. F., West, A. H., Stock, J. B., Ringe, D., and Petsko, G. A. (1993) Structure of the Mg^{2+}-Bound Form of CheY and Mechanism of Phosphoryl Transfer in Bacterial Chemotaxis, *Biochemistry 32*, 13375-13380.

17. Santoro, J., Bruix, M., Pascual, J., Lopez, E., Serrano, L., and Rico, M. (1995) Three-Dimensional Structure of Chemotactic CheY Protein in Aqueous Solution by Nuclear Magnetic Resonance Methods, *J. Mol. Biol. 247*, 717-725.

18. Lee, S.-Y., Cho, H. S., Pelton, J. G., Yan, D., Berry, E. A., and Wemmer, D. E. (2001) Crystal Structure of Activated CheY. Comparison With Other Activated Receiver Domains, *J. Biol. Chem. 276*, 16425-16431.

19. Simonovic, M., and Volz, K. (2001) A Distinct Meta-active Conformation in the 1.1-Å Resolution Structure of Wild-type ApoCheY, *J. Biol. Chem. 276*, 28637-28640.

20. Lee, S.-Y., Cho, H. S., Pelton, J. G., Yan, D., Henderson, R. K., King, D. S., Huang, L.-S., Kustu, S., Berry, E. A., and Wemmer, D. E. (2001) Crystal structure of an activated response regulator bound to its target, *Nat. Struct. Biol. 8*, 52-56.

21. Bourret, R. B., Drake, S. K., Chervitz, S. A., Simon, M. I., and Falke, J. J. (1993) Activation of the Phosphosignaling Protein CheY. II. Analysis of Activated Mutants by 19F NMR and Protein Engineering, *J. Biol. Chem. 268*, 13089-13096.

22. Lukat, G. S., Lee, B. H., Mottonen, J. M., Stock, A. M., and Stock, J. B. (1991) Roles of the Highly Conserved Aspartate and Lysine Residues in the Response Regulator of Bacterial Chemotaxis, *J. Biol. Chem. 266*, 8348-8354.

23. Ganguli, S., Wang, H., Matsumura, P., and Volz, K. (1995) Uncoupled Phosphorylation and Activation in Bacterial Chemotaxis. The 2.1-Å Structure of a Threonine to Isoleucine Mutant at Position 87 of CheY, *J. Biol. Chem. 270*, 17386-17393.

24. Zhu, X., Amsler, C. D., Volz, K., and Matsumura, P. (1996) Tyrosine 106 of CheY Plays an Important Role in Chemotaxis Signal Transduction in *Escherichia coli*, *J. Bacteriol. 178*, 4208-4215.

25. Sanders, D. A., Gillece-Castro, B. L., Stock, A. M., Burlingame, A. L., and Koshland, D. E., Jr. (1989) Identification of the Site of Phosphorylation of the Chemotaxis Response Regulator Protein, CheY, *J. Biol. Chem. 264*, 21770-21778.

26. Welch, M., Oosawa, K., Aizawa, S.-I., and Eisenbach, M. (1994) Effects of Phosphorylation, Mg^{2+}, and Conformation of the Chemotaxis Protein CheY on Its Binding to the Flagellar Switch Protein FliM, *Biochemistry 33*, 10470-10476.

27. Bourret, R. B., Hess, J. F., and Simon, M. I. (1990) Conserved aspartate residues and phosphorylation in signal transduction by the chemotaxis protein CheY, *Proc. Natl. Acad. Sci. U. S. A. 87*, 41-45.

28. Appleby, J. L., and Bourret, R. B. (1998) Proposed Signal Transduction Role for Conserved CheY Residue Thr87, a Member of the Response Regulator Active-Site Quintet, *J. Bacteriol. 180*, 3563-3569.

29. Roman, S. J., Meyers, M., Volz, K., and Matsumura, P. (1992) A Chemotactic Signaling Surface on CheY Defined by Suppressors of Flagellar Switch Mutations, *J. Bacteriol. 174*, 6247-6255.

30. Shukla, D., Zhu, X. Y., and Matsumura, P. (1998) Flagellar Motor-switch Binding Face of CheY and the Biochemical Basis of Suppression by CheY Mutants That Compensate for Motor-switch Defects in *Escherichia coli*, *J. Biol. Chem. 273*, 23993-23999.

31. Bren, A., and Eisenbach, M. (1998) The N Terminus of the Flagellar Switch Protein, FliM, is the Binding Domain for the Chemotactic Response Regulator, CheY, *J. Mol. Biol. 278*, 507-514.

32. Blat, Y., and Eisenbach, M. (1996) Conserved C-Terminus of the Phosphatase CheZ Is a Binding Domain for the Chemotactic Response Regulator CheY, *Biochemistry 35*, 5679-5683.

33. Silversmith, R. E., Smith, J. G., Guanga, G. P., Les, J. T., and Bourret, R. B. (2001) Alteration of a Nonconserved Active Site Residue in the Chemotaxis Response Regulator CheY Affects Phosphorylation and Interaction with CheZ, *J. Biol. Chem. 276*, 18478-18484.

34. Welch, M., Chinardet, N., Mourey, L., Birck, C., and Samama, J.-P. (1998) Structure of the CheY-binding domain of histidine kinase CheA in complex with CheY, *Nat. Struct. Biol. 5*, 25-29.

35. Zhu, X., Volz, K., and Matsumura, P. (1997b) The CheZ-binding Surface of CheY Overlaps the CheA- and FliM-binding Surfaces, *J. Biol. Chem. 272*, 23758-23764.

36. Bilwes, A. M., Alex, L. A., Crane, B. R., and Simon, M. I. (1999) Structure of CheA, a Signal-Transducing Histidine Kinase, *Cell 96*, 131-141.

37. McEvoy, M. M., Hausrath, A. C., Randolph, G. B., Remington, S. J., and Dahlquist, F. W. (1998) Two binding modes reveal flexibility in kinase/response regulator interactions in the bacterial chemotaxis pathway, *Proc. Natl. Acad. Sci. U. S. A. 95*, 7333-7338.

38. Park, S.-Y., Beel, B. D., Simon, M. I., Bilwes, A. M., and Crane, B. R. (2004) In different organisms, the mode of interaction between two signaling proteins is not necessarily conserved, *Proc. Natl. Acad. Sci. U. S. A. 101*, 11646-11651.

39. Kirby, J. R., Kristich, C. J., Saulmon, M. M., Zimmer, M. A., Garrity, L. F., Zhulin, I. B., and Ordal, G. W. (2001) CheC is related to the family of flagellar switch proteins and acts independently from CheD to control chemotaxis in Bacillus subtilis, *Mol. Microbiol. 42*, 573-585.

40. Guhaniyogi, J., Robinson, V., and Stock, A. (2006) Crystal Structures of Beryllium Fluoride-free and Beryllium Fluoride-bound CheY in Complex with the Conserved C-terminal Peptide of CheZ Reveal Dual Binding Modes Specific to CheY Conformation, *J. Mol. Biol. 359*, 624-645.

41. Halkides, C. J., Zhu, X., Phillion, D. P., Matsumura, P., and Dahlquist, F. W. (1998) Synthesis and Biochemical Characterization of an Analogue of CheY-Phosphate, a Signal Transduction Protein in Bacterial Chemotaxis, *Biochemistry 37*, 13674-13680.

42. McAdams, K., Mesecar, A., Santarsiero, B., Casper, E. S., Haas, R. M., and Halkides, C. J. (2006) The Structures of T87I Phosphono-CheY and T87I/Y106W Phosphono-CheY and Their Binding Affinities to the FliM and CheZ Peptides, *Biochemistry*, Submitted for Review.

43. Park, S.-Y., Lowder, B., Bilwes, A. M., Blair, D. F., and Crane, B. R. (2006) Structure of FliM provides insight into assembly of the switch complex in the bacterial flagella motor, *Proc. Natl. Acad. Sci. U. S. A. 103*, 11886-11891.

44. Alpert, A. J. (1988) Hydrophobic interaction chromatography of peptides as an alternative to reversed-phase chromatography, *J. Chromatog. A 444*, 269-274.

45. Alpert, A. J., and Gilbert, H. F. (1985) Detection of oxidized and reduced glutathione with a recycling postcolumn reaction, *Anal. Biochem. 144*, 553-562.

46. Gill, S. C., and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data, *Anal. Biochem. 182*, 319-326.

47. Ellman, G. L. (1959) Tissue sulfhydryl groups, *Arch. Biochem. Biophys. 82*, 70-77.

48. McAdams, K. (2002) Structural and Binding Studies of An Active State Analog of T87I/Y106W CheY, in *Department of Chemistry*, p 83, University of North Carolina Wilmington, Wilmington.

49. Scopes, R. K. (1987) *Protein Purification*, Second ed., Springer-Verlag, New York.

50. Hirs, C. H. W. (1977) *Gel Filtration*, Vol. 47, Academic Press, New York.

51. Jocelyn, P. C., and William, B. J. a. O. W. G. (1987) Chemical reduction of disulfides, in *Methods Enzymol.*, pp 246-256, Academic Press.

52. Scopes, R. K. (1974) Measurement of protein by spectrophotometry at 205 nm, *Anal. Biochem. 59*, 277-282.

53. van Iersel, J., Jzn, J. F., and Duine, J. A. (1985) Determination of absorption coefficients of purified proteins by conventional ultraviolet spectrophotometry and chromatography combined with multiwavelength detection, *Anal. Biochem. 151*, 196-204.

54. Houghten, R. A., and Li, C. H. (1979) Reduction of sulfoxides in peptides and proteins, *Anal. Biochem. 98*, 36-46.

55. Guggenheim, E. A. (1926) *Philisophical Magazine 2*, 538.

56. Mant, C. T., and Hodges, R. S. (1991) *HPLC Terminology: Practical and Theoretical*, CRC Press LLC, Boca Raton, FL.

57. Nordlund, H. R., Laitinen, O. H., Uotila, S. T. H., Nyholm, T., Hytonen, V. P., Slotte, J. P., and Kulomaa, M. S. (2003) Enhancing the Thermal Stability of Avidin. Introduction of Disulfide Bridges between subunit interfaces, *J. Biol. Chem. 278*, 2479-2483.

58. Deutschman, W. A., and Dahlquist, F. W. (2001) Thermodynamic Basis for the Increased Thermostability of CheY from the Hyperthermophile Thermotoga maritima, *Biochemistry 40*, 13107-13113.